THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XVIII | May, 1927 | Number 5

EQUALLY OFTEN NOTICED DIFFERENCES[1]

L. L. THURSTONE

University of Chicago and Institute for Juvenile Research

Quality scales such as Thorndike's scales for handwriting and for children's drawings are constructed on the principle that equally often noticed differences are equal. It is the object of this paper to examine this principle of scale construction and to show that the principle is valid only when the stimuli or specimens are uniform in the degree of ambiguity or dispersion of excellence that is perceived in them.

The principle that equally often noticed differences are equal is a psychological commonplace. Of course the principle does not say anything whatever about the physical differences between the stimuli. That is a totally different matter. Some stimuli are of such a character that a physical scale can be postulated, parallel to the psychological scale. An example is the lifted weight series. The weights may be allocated on a physical scale in grams and the same weights may be allocated on a psychological scale in accordance with the principle that equally often noticed differences are equal. The construction of an educational quality scale is not ordinarily concerned with the physical scale.

For some kinds of stimuli there exists no physical scale that can be at all readily defined or measured. The excellence of handwriting or of freehand drawing is an example in which the specimens may be allocated on a psychological scale of excellence, as perceived, but for which it is at least difficult to define a parallel physical scale. However, it is possible to construct a quality scale without giving any consideration to the possibility of a parallel physical scale. It is an interesting inquiry to search for the physical measurements that may parallel the psychological measurements but the present analysis is concerned

[1] This is one of a series of articles prepared by members of the staff of Illinois Institute for Juvenile Research, Chicago. Series B105.

only with the psychological scale which is constructed on the basis of relative frequency of discrimination between the specimens.

When we say that equally often noticed differences are psychologically equal, it sounds so obvious that we are tempted to regard the assertion as a psychological axiom. But the matter is far from being

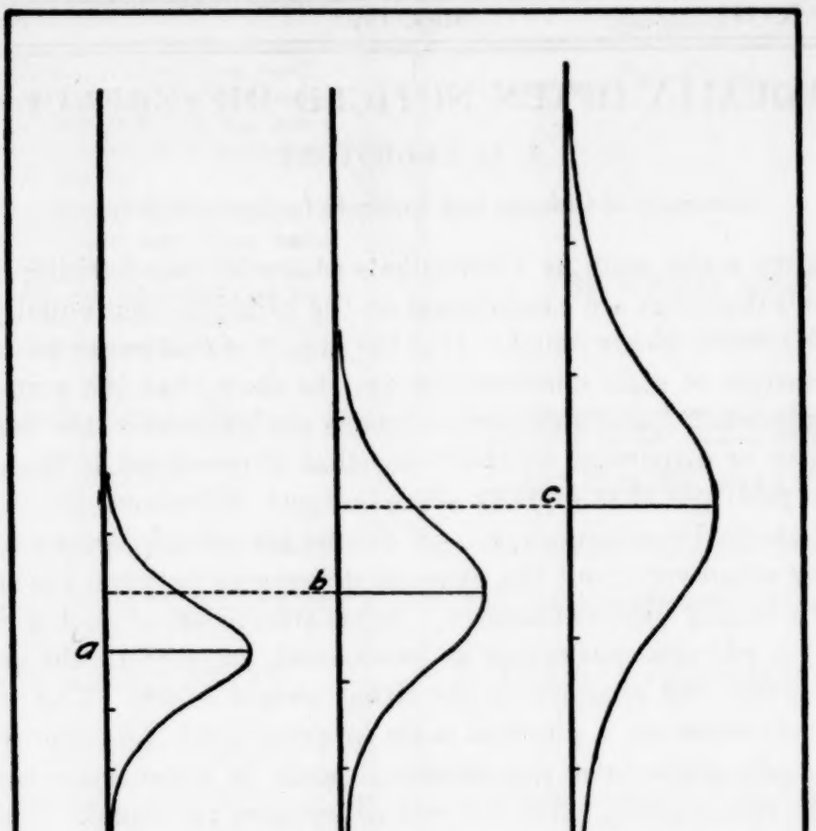


FIG. 1.

so simple. In fact, *equally often noticed differences are not always equal, not even psychologically.* The matter can be made clear by considering a simple case.

In Fig. 1 let the vertical dimension represent a psychological continuum such as Thorndike's handwriting scale. Let *a* represent the scale value of one of the handwriting specimens *A*. The small probability curve may represent the fluctuation in scale value as

perceived by several hundred judges. Let the spread or dispersion of this distribution be designated σ_a. Let b and c represent the scale values of two other handwriting specimens B and C. Suppose that specimen C is more ambiguous to judge for degree of excellence than B, and that B is more ambiguous than A. This means, merely, that the several hundred judges differ considerably in the degree of excellence that they see in specimen C and that they agree more closely as to the excellence they perceive in B while specimen A shows the best agreement with the smallest dispersion in perceived excellence. In Fig. 1 the three distributions are so drawn that

$$\sigma_a = \tfrac{2}{3}\sigma_b$$

and

$$\sigma_b = \tfrac{2}{3}\sigma_c$$

The scale value b is the mean degree of perceived excellence in specimen B and it is represented at the level $+1\sigma$ on the distribution for specimen A. In exactly the same proportional arrangement, the mean perceived excellence of C, its scale value, is located at the level of $+1\sigma$ on the distribution for specimen B.

It is readily seen in Fig. 1 that the overlapping or confusion between A and B is the same as that between B and C. Hence we should expect the same proportion of judgments "B better than A" as of judgments "C better than B." By the application of the law of comparative judgment[1] to the present case, the proportion

$$p_{B>A} = p_{C>A} = .71$$

We have here a situation in which the two scale differences $(B - A)$ and $(C - B)$ which are clearly unequal on the psychological scale are nevertheless equally often noticed. In fact

$$(C - B) = \tfrac{3}{2}(B - A).$$

Without considering the validity of the fundamental psychophysical equation by which the exact proportion .71 is determined, it is clear from inspection of Fig. 1 that the proportion of judgments "C better than B" must be equal to the proportion "B better than A" because the relations between A and B are duplicated in the relations between B and C.

We conclude, therefore, that equally often noticed differences are equal only when the discriminal dispersions of the specimens are equal.[2]

[1] Thurstone: "A Law of Comparative Judgment." To be published in the *Psychologicrl Review.*

[2] For a further statement of the concept of discriminal dispersion, see my article, Psychophysical Analysis. To be published in *The American Journal of Psychology.*

Figure 2 has been drawn to show the converse statement, namely, that equal psychological separations are not necessarily equally often noticed. Let Fig. 2 represent three specimens, A, B, C. Let the specimens A and B be easily perceived with very little ambiguity while specimen C is difficult to evaluate. Specimen C is therefore shown with a large dispersion as compared with the dispersions of A and B.

Since B and C have the same scale value it follows of course that

$$(B - A) = (C - A)$$

Now consider what will happen when the judges are asked to compare specimens A and B. Since the lowest excellence that any judge sees in B is approximately equal to the highest excellence that any judge sees in A, we should expect a very high proportion of judgments "B better than A," perhaps 90 or 95 per cent.

The judges are then asked to compare B and C. Half of them see specimen B as more excellent than the scale value b, and half of them see specimen C as more excellent than the scale value c. If we assume for the sake of simplicity that the variations in perceived excellence of specimens A, B, C, are uncorrelated, it is evident that the percentage of judgments "B better than C" will be .50. Hence these two specimens B and C are declared to have the same scale value and they are so indicated in Fig. 2.

Now let the judges compare specimens A and C. Notice that there is considerable overlapping in judgments of excellence for A and C. For example, a considerable proportion of the degrees of perceived excellence for C is below the average judgment for A. Hence we shall expect to find a proportion of judgments "C better than A" about 70 or 80 per cent but certainly not 90 or 95 per cent.

Since $B = C$ it follows that

$$B - A = C - A$$

and that the difference $(B - A)$ should be as often discriminated as $(C - A)$ but we have seen in Fig. 2 that if the stimuli or specimens are unequal in relative degree of ambiguity the two equal differences $(B - A)$ and $(C - A)$ will not be equally often noticed. We conclude therefore that *equal psychological differences are equally often noticed only when the discriminal dispersions of the specimens are equal.*

If the discriminal dispersions can be assumed to be constant for all of the specimens in a series or scale, then the principle that equally

often noticed differences are equal is valid, but it is hardly likely that the assumption is always justified. In psychophysical experimentation with lifted weights, for example, it is probably safe to assume that the ambiguity is constant because the weight cylinders are all alike, in appearance, and they are even kept at the same temperature by careful experimenters so as to avoid thermal cues caused by having the sub-

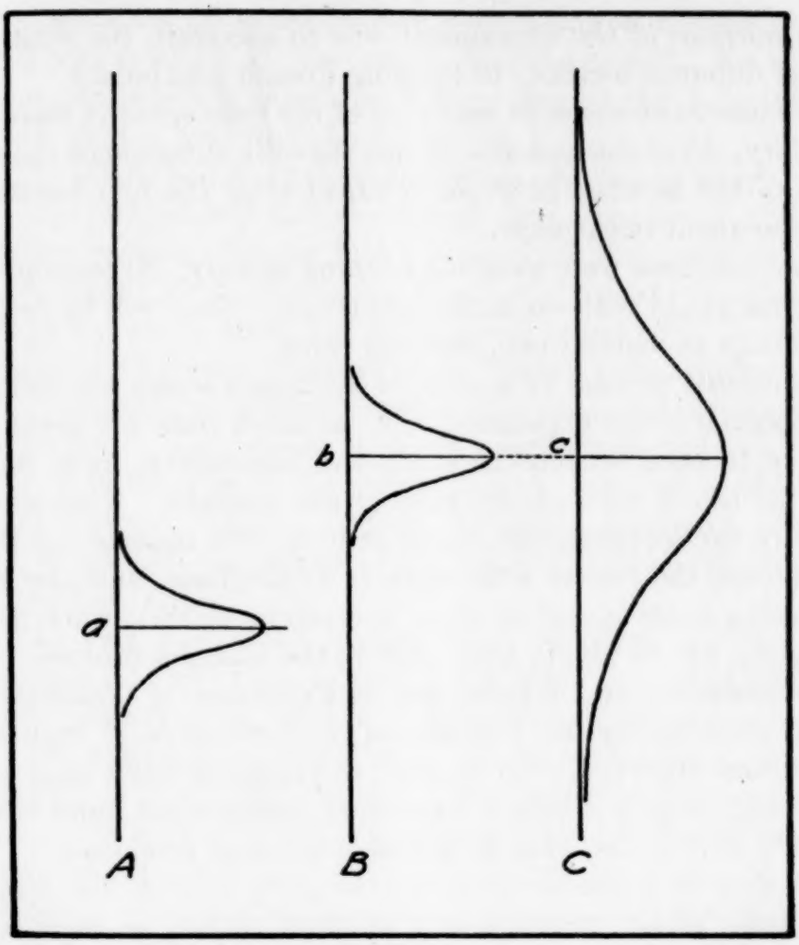


FIG. 2.

ject handle the standard weight-cylinder more frequently than the comparison-cylinders. But it is hardly likely that specimens of English composition, for example, are sufficiently homogeneous in ambiguity of excellence to justify the assumption that equally often noticed differences are equal. Certainly this is an assumption to be experimentally verified in each case. Equally often noticed differences are not necessarily equal, not even subjectively or psychologically.

AN EXPERIMENT IN LEARNING FRENCH VOCABULARY

LOUISE C. SEIBERT

Goucher College

Purpose

The purpose of the experiment was to ascertain the relative efficiency of different methods of learning French vocabulary.

The measurements were made (*a*) of the time spent in learning the vocabulary, (*b*) of the amount of time saved in subsequent relearnings and (*c*) of the number of words retained after the first learning and after subsequent relearnings.

Three methods were used (*a*) learning silently, (*b*) learning aloud, (*c*) learning aloud with an immediate recall. They will be designated in this study as method one, two and three.

The normal process of a student learning a lesson was followed as far as possible in the experiment; *i.e.*, as much time was given as was necessary to each student to study the vocabulary, until she could repeat the list of words twice without any mistake. This, of course, was really overlearning, but it was done for two reasons: for the sake of comparing the results with those of Ebbinghaus, and also because overlearning is advocated by every competent teacher, when mechanical memory has to play a large part in the learning process. This is what Norsworthy and Whitley say in *Psychology of Childhood:* "The value of overlearning, for purposes of recall needs to be impressed on teachers and students. To be able to repeat a thing once without error though it may satisfy a laboratory requirement, does not argue a memory of it in the sense of probable accurate retention. The correctness may be a matter of chance, as every learner discovers when 'trying once more'; consequently children should be encouraged to learn till they can repeat material at least twice running without error, which will entail a much greater number of repetitions and efforts to recall."

Experiments somewhat similar have been undertaken, the foreign words being replaced by non-sense syllables by Ebbinghaus, but the results obtained could hardly be applied to the study of foreign words, for the association of a foreign word with its meaning plays a great part in the memorizing process. Thorndike says: "The common inference from Ebbinghaus's figures that half the effect of memorizing

is lost in less than an hour and two-thirds in a day is very far from holding true of memory for paired associates. This inference was found by Professor Cattell to be false also in the case of vocabularies learned as series."

The Experiment

Choice of Vocabulary.—While Ebbinghaus attempted to secure unknown vocables of equal difficulty by using non-sense syllables, he and others questioned whether he had succeeded. To solve the more complex problem involving foreign words, the following principles were used in selecting the vocabulary:

(*a*) The words had to be absolutely unknown to the group. There was no way to ascertain this absolutely until the time of the experiment, when it would be too late. But, since the troup tested was known (*cf. infra*) we could rely upon our experience of the vocabulary likely to be known by them and select words outside it. In this selection, technical, obsolete or rare words in English were avoided, as their possible unfamiliarity to the student would have vitiated the value of the experiment. Also words of Latin origin were omitted, as their similarity with the French derivative throws them out of the class of the usual foreign vocable. These conditions restricted singularly the field of possible words; however, a list of about 400 words was established tentatively.

(*b*) The lists had to be of the same difficulty. From the list, eight series had to be made, in such a way that they could fairly compare with each other, in respect to difficulty. This was not an easy task, for we came immediately upon the question: What constitutes the difficulty of a foreign vocable? The length of the word? The difficulty of its spelling? Of its pronounciation? The greater or less familiarity of its meaning? The lack of the possibility of complex associations in the mind of the student learning it at the time? The place it may happen to occupy on the list? The problem we think has not yet been solved. The relative difficulty of a given list of words can be ascertained by means of experiments, but we have no principles to guide us, as far as the difficulty of the words outside the list is concerned. It is even a question whether a word has any intrinsic difficulty. It may be that this difficulty depends almost entirely on the individual, on the associations a given word suggests at one time and not at another, etc.

It has been recognized that sometimes the most "difficult" words are the ones which are remembered better, and also that what is diffi-

cult for one person is not necessarily difficult for another. To attempt to equalize the difficulty as far as possible, we selected words similar in the following respect:

1. The meaning in English was perfectly familiar to the group tested.
2. Each of the eight lists contained the same number of words of one syllable, of two, of three and four syllables. (There were no words of more than four syllables.) The total number of syllables were therefore the same in each list. Moreover, there was a correspondence in the words as to the makeup of the syllables, a similar number in each list being composed of a consonant and a vowel, and another similar number composed of consonant and diphthong. The relative position of these matched French words, did not correspond on each list, however, when the test was given.
3. The number of syllables of the corresponding English words, were in the majority of cases the same as those of the French ones and, when this did not occur, at least the number of syllables of the English correspondents were the same.

However, in spite of the care taken, the lists could not be assumed to be of the same difficulty. Therefore, in order to counterbalance the difference, the lists were rotated. The same list would be given to one group to be studied silently, while it would be given to another group to be studied aloud and to a third to be studied aloud with a written recall. This practice was carefully observed throughout the experiment, so that the results obtained were, it is felt, fairly independent of the degree of difficulty of one particular list.

The students were asked to mark, in a specified way, any word they had seen before even if they had the faintest recollection of it, and any list thus marked was discarded when the results were tabulated. As a matter of fact, only two such cases occurred and these happened when the test was given to a group for which it had not been primarily devised.

The Group Tested.—The 81 students for whom we have completed records were members of three sections which were rather homogeneous in respect to training. They had had the equivalent of one year of college work in French.

Rotations of Lists and Tests.—The experiment was divided into three parts.

1. A series was to be learned silently, then relearned after an interval of two days, ten days, and forty-two days.
2. Another series was learned aloud, then relearned after an interval of two days, ten days, and forty-two days.
3. Another series was learned and relearned at the same intervals, but this time aloud, with an immediate written recall.

To compensate for the possible effects of practice, the tests were alternated, that is, while one group began the series to be learned silently, another began the series to be learned aloud and the third began the series to be learned with a written recall.

The Giving of the Test.—At the beginning of the class period, each student was given a card on which the list of words was written. The English words were given first, then the French translation. The cards were turned upside down, so that no words could be seen. The students were also given a slip of paper, on which they wrote their names, the section, the date, leaving a space to record the time and, later on, their score. On the other side of the slip, they were to write the words they could remember.

The students were told to study the list in the manner indicated, until they could repeat the words twice without a mistake. (The English words of the list had been written on the board in a different order and it was this list that the student used to control the learning of the words.) If a mistake was made they were told to begin again. As soon as a student was able to repeat the list twice, she recorded the time on the slip in the space provided for that purpose, laid the slip aside and began the class work.[1]

Recording Time.—The teacher recorded the time at the board. At the beginning of a minute, 0 was marked and the student started. The time was recorded every 15 seconds. When a student finished she recorded the time last written on the board.

The Recalls. First Recall.—At the end of the class period the English words were written on the board in a different order and the students were told to write on the back of their slip on which the time had been recorded as many words as they could remember. The English word could be omitted if the corresponding French one was given with its right number. The papers were then collected.

Second Recall and First Relearning.—Two days later at the beginning of the class period, the same card and a slip of paper were given to the student. The English words of the list were written on the board in a still different order, and the students wrote as many words as they could remember with the customary indications of date, section, etc., on the slip. A place was left for recording the time of relearning. The students were allowed a reasonable amount of

[1] In the third method the time necessary for writing once the list of words was not counted.

time for the procedure. Then they relearned the list using the same procedure as in the original learning.

The same procedure was observed for the second and third recalls at intervals of 10 and 42 days:

The Scoring of the Tests.—A word given correctly was scored 3; if there was a mistake in the accents the word was scored 2; if a letter was missing or wrong, the word was scored 1. For all other errors, the word was scored 0. An individual's score was the sum of the scores on all the words of the list.

Tabulating the Results

A list of all the students having taken the tests was made and before each name was recorded the time and score for the three tests and their recalls. Any student who had missed either a test or a recall was eliminated, so as to have an identical group through all the series of tests. The results were then tabulated in different ways. Frequency tables for the distributions of time and scores were made, both for the tests and the recalls. The means and SD of the scores for each method of learning are shown in Table I.

Table I.—Time (in Quarter Minutes) of Learning and Relearning after Certain Intervals

	First learning		First relearning 2 days		Second relearning 10 days		Third relearning 42 days	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Studying silently.....	20.01	6.0	6.2	3.1	4.9	3.5	5.9	2.5
Studying aloud.....	20.01	7.2	5.2	3.0	3.9	2.1	5.2	2.3
Studying aloud (written recall).....	17.0	4.8	5.2	2.0	4.0	2.1	5.7	2.0

Table II.—Accuracy of Learning and Relearnings after Certain Intervals

	After 50 minutes		After 2 days (No relearning)		After 10 days (First relearning)		After 42 days (Second relearning)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Studying silently.....	27.1	8.0	22.7	8.9	23.0	9.0	16.6	9.7
Studying aloud.....	29.5	6.6	26.5	7.8	26.3	7.1	18.3	8.9
Studying aloud (written recall).....	29.7	7.4	23.4	8.4	23.3	9.2	14.9	7.4

Graphical Representations.—In the following graphical representations, the time unit is the fourth of a minute and the scoring unit is one.

Time Distributions.—The distributions have been arranged so that a constant comparison between the performance in the three different processes can easily be made. Figure 1 (*a*, *b*, *c*) represents respectively the time taken to learn the vocabulary in the three different ways indicated.

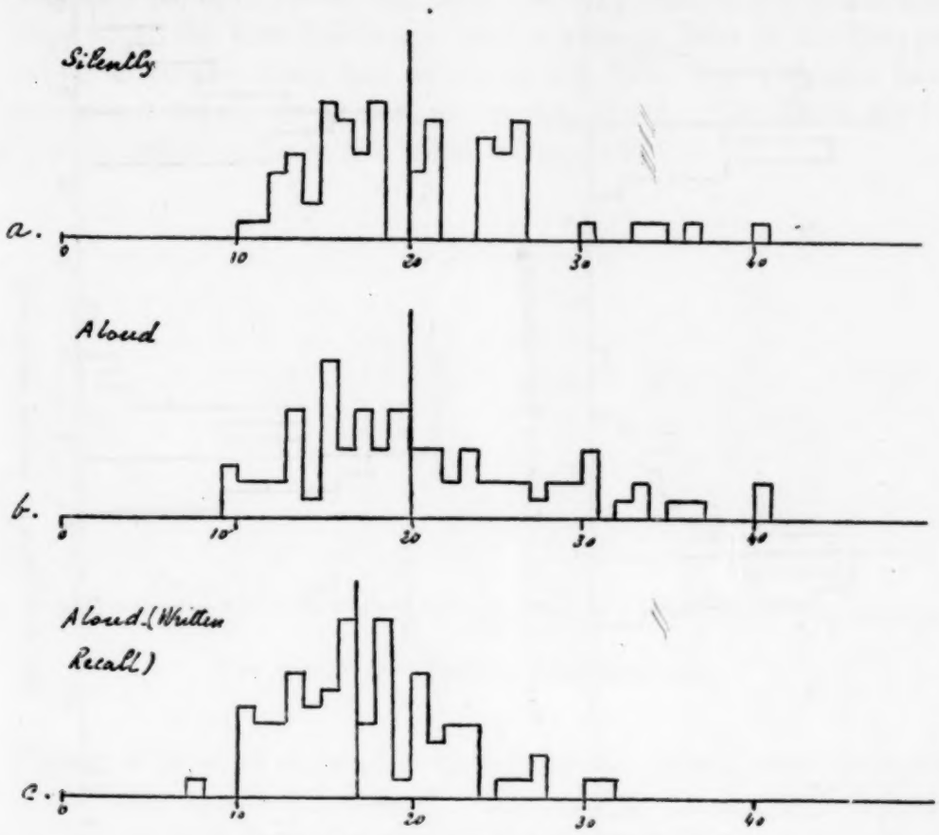


FIG. 1 (*a*, *b*, *c*).—Time for learning.

It will be noticed at a glance that the distribution in Fig. 1(*a*) is more widely spread than in (*b*) and (*c*). While the average in (*b*) is the same as in (*a*) the SD is 7.2 against 6.0, There is a conspicuous lagging group. The comparison of the averages is also interesting. While the average between (*a*) and (*b*) is almost identical, the average for (*c*) is 17, a marked improvement. But in the learning process (*b*) and (*c*) were identically alike, since the time for the written recall was not recorded. Consequently we should have expected a somewhat similar result. The difference between the two averages was found to

be due to a lagging group. Group 3 had an average of 25.5 against 19.0 and 18.0 for groups 2 and 1. This group took the test late in the afternoon.

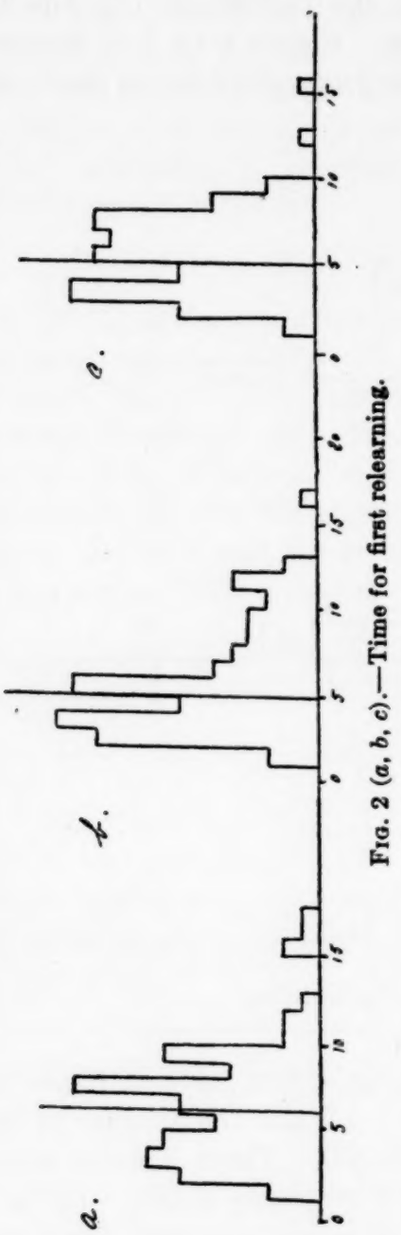


FIG. 2 (*a*, *b*, *c*).—Time for first relearning.

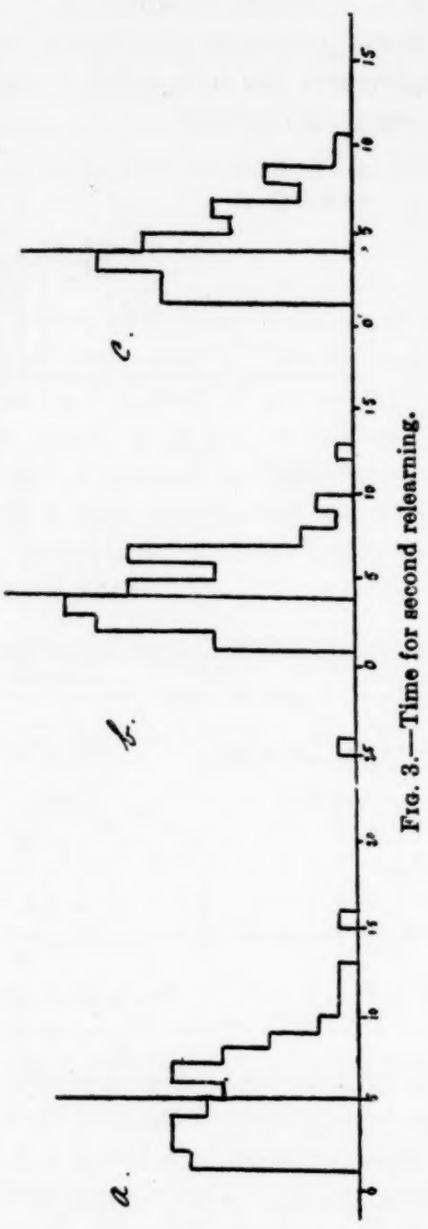


FIG. 3.—Time for second relearning.

The results may be roughly recapitulated as follows: While there is no perceptible gain of time in studying aloud instead of studying silently, there is a remarkable improvement in the third process, both in the range of the distribution and the average.

Figure 2 (*a*, *b*, *c*,) represents the time taken to relearn the vocabulary in the three different ways. Compared with each other we find that there is a decided gain in the amount of time saved in relearning by the two last processes, the average being 5.2 for both of them against 6.2 for the first. There is a more marked compactness in (*c*) where the SD is 2.0 against 3.1 and 3.0 respectively for the others.

Figure 3 (*a*, *b*, *c*) shows the time taken to relearn the vocabulary 10 days after the first learning. The saving of time in the two last processes is greater than the saving in the first, the averages being respectively 3.9 and 4.0 against an average of 4.9. The SD in the two last distributions is the same, being 2.1 against 3.5.

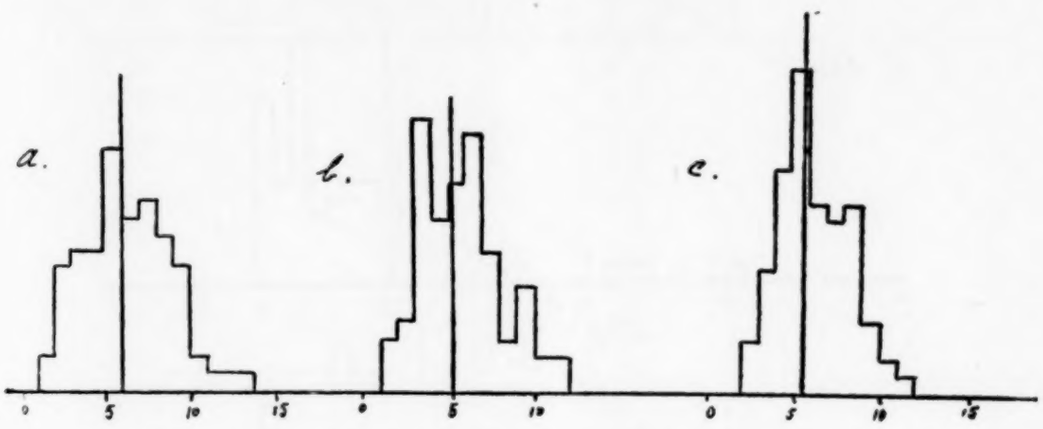


FIG. 4 (*a*, *b*, *c*).—Time for third relearning.

Figure 4 (*a*, *b*, *c*) shows the time taken to relearn the vocabulary 42 days after the first learning. The greater saving is in learning aloud, the average being 5.2 against 5.9 for the relearning silently and 5.7 for relearning aloud with a written recall, although the written recall gives the most compact distribution.

Recapitulation.—It appears then, that the question as to whether it takes less time to learn by one process or the other is still unsettled. The chances are, however, that there is a gain of time in studying aloud, since there is a definite gain in one case and very little difference in the other.

As for the saving of time in the relearnings, the two last methods are decidedly superior with a slight advantage of the second over the third.

It must be noticed also that all the time distributions are rather compact, and that the skewing of the curves is practically the same.

Accuracy Distributions.—Figure 5 (*a*, *b*, *c*) represents the distributions of accuracy, 50 minutes after learning. (*a*) (*b*) and (*c*) are all widely spread and skewed similarly.

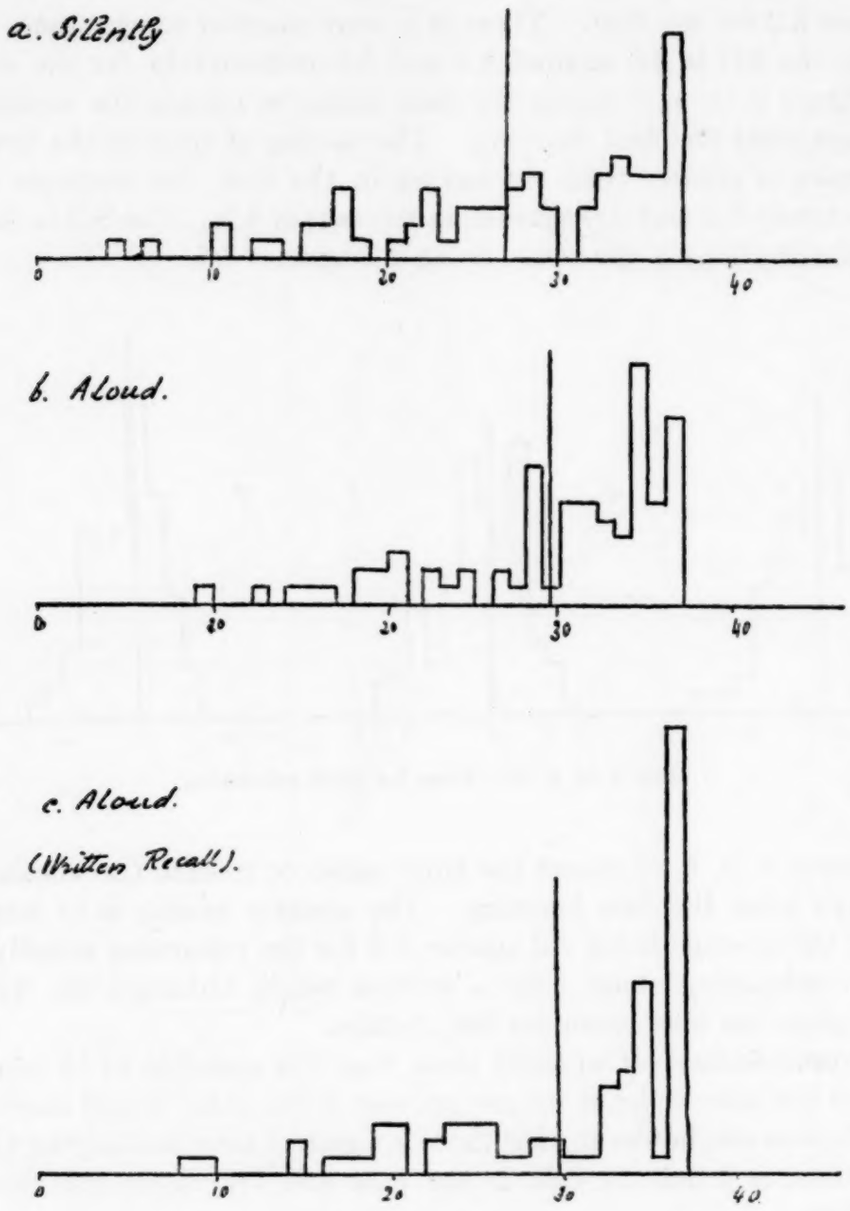


FIG. 5 (*a*, *b*, *c*).—Accuracy after first learning.

The third method gives the best results and the first the poorest. The third method also yields the most perfect scores.

Figure 6 (*a*, *b*, *c*) presents the accuracy after an interval of two days. The distributions continue to be spread widely, although it is more compact in (*b*), which shows also a better average.

Figure 7 (*a*, *b*, *c*) represents the accuracy after an interval of 10 days. The distributions are again widely spread and again the most compact one is in (*b*), which also shows the best average.

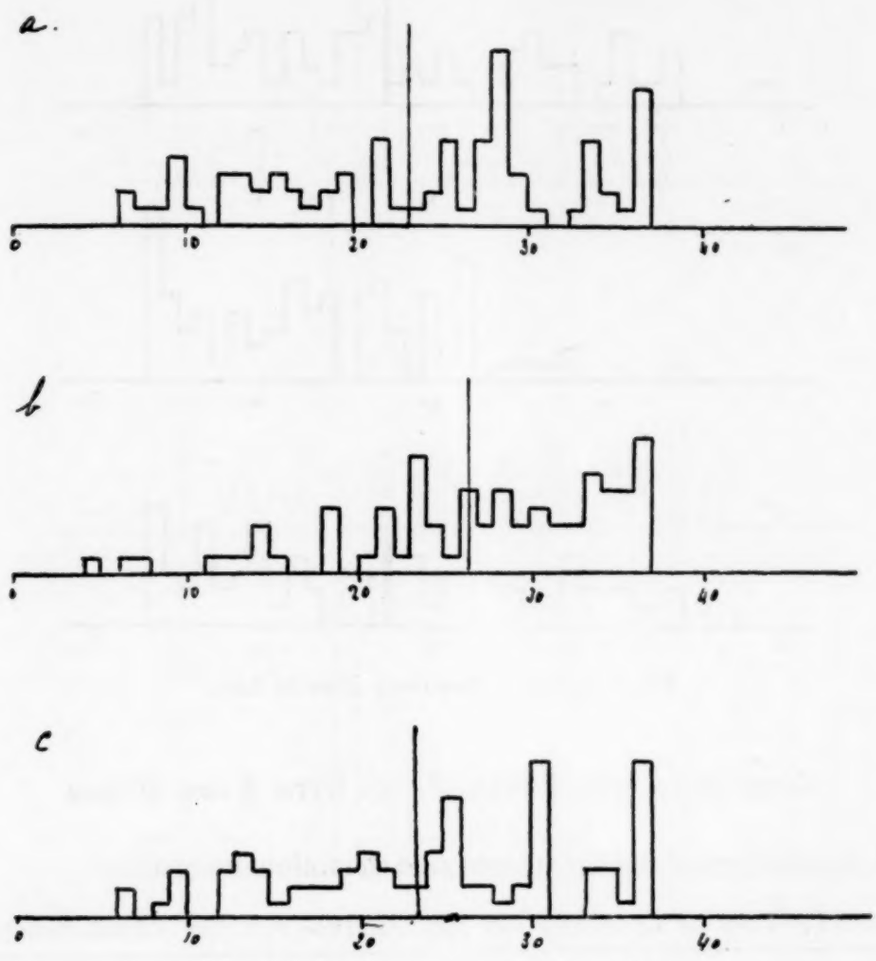


FIG. 6 (*a*, *b*, *c*).—Accuracy after 10 days.

Figure 8 (*a*, *b*, *c*) represents the accuracy after an interval of 42 days. The average of (*b*) is the highest. Let us note here that the showing of the third process is poorer than the two other processes. The distributions also vary in character, (*b*) being almost rectangular, while (*a*) is very irregular and (*c*) is skewed very markedly.

Recapitulation.—In all cases the best averages for accuracy are made when the words have been learned aloud. The showings or the third process are as poor as those of the first and sometimes worse, and there is a striking similarity between the two curves.

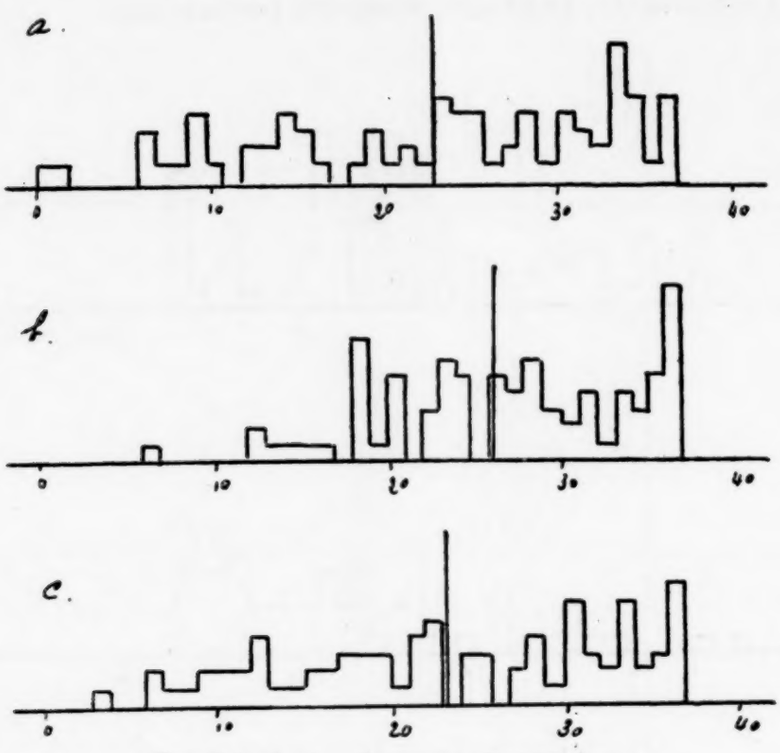


FIG. 7 (*a*, *b*, *c*).—Accuracy after 10 days.

COMPARING THE THREE TESTS WITH EACH OTHER

A comparison of the three tests give the following results:

TABLE III.—TIME OF LEARNING AND RELEARNINGS FOR THE THREE PROCESSES

	First learning, units	First relearning, units	Second relearning, units	Third relearning, units
Studying silently.....	1621	499	401	480
Studying aloud.....	1632	422	318	425
Studying aloud (written recall).....	1376	422	324	465

1. Studying aloud took .67 per cent[1] more time than studying silently, but in the third process studying aloud took 15.1 per cent less time than studying silently.

2. Relearning aloud took 15.4 per cent less time than relearning silently, and relearning with a written recall took also 15.4 per cent. less time than the first process.

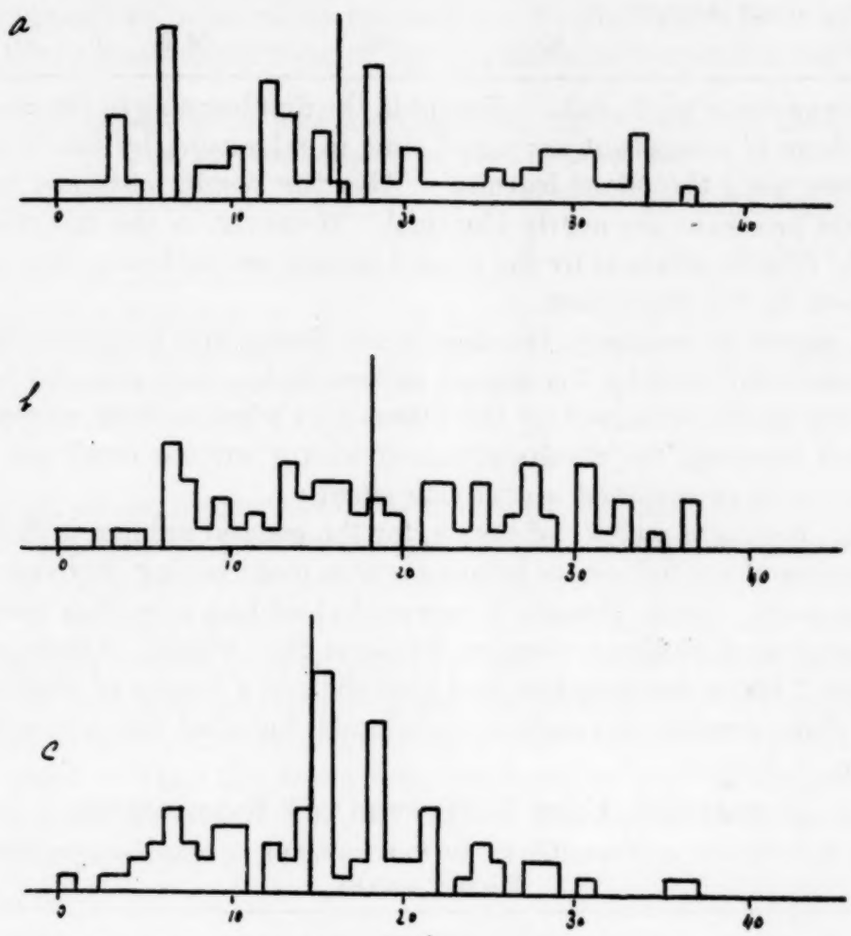


FIG. 8 (*a*, *b*, *c*).—Accuracy after 42 days.

3. The second process took 20.7 per cent less time than the first, and the third process took 19.2 per cent less time than the first.

4. The second process took 11.4 per cent less time than the first, and the third process took 3.8 per cent less time than the first.

---

[1] The percentages are all calculated on the time taken to study silently.

TABLE IV.—ACCURACY OF LEARNING AND RELEARNINGS FOR THE THREE PROCESSES, IN PERCENTAGES OF THE PERFECT SCORE

	First learning	First relearning	Second relearning	Third relearning
Studying silently.....	75.3	63.1	64.0	46.3
Studying aloud.....	81.9	73.6	73.3	51.0
Studying aloud (written recall).....	82.5	65.1	64.9	41.5

Discussion of the Results.—Except in the first learning of the vocabulary, there is a constant saving of time in relearning by the two last processes over the silent learning. Also the results obtained in the two last processes are nearly identical. However, in the last relearning, the results obtained by the second process are far better than those obtained by the two others.

In regard to accuracy, the results are clearer and very interesting. The result obtained by the second process is, in every case, far better than the results obtained by the others and what is more, except for the first learning, the results obtained with a written recall are very similar to those obtained in studying silently.

This finding is worthy of notice, for the general opinion is that the impression of several senses brings about a more lasting impression on the memory. Here, though, it seems that adding a written recall to the visual and auditory memory weakens the results. According to Gordon, "Other investigators find that there is a danger of distraction when three sensory elements are combined, but that two are superior to one."[1]

COMPARING EACH TEST WITH THE RELEARNINGS

TABLE V.—SAVING OF TIME (GIVEN IN PERCENTAGE) IN THE RELEARNINGS FOR EACH PROCESS

	Silently	Aloud	Aloud (written recall)
Time taken for first learning	1621 Units (t1)	1632 Units (t2)	1376 Units (t3)
Time saved in first relearning.....	69.2 of t1	74.1 of t2	69.4 of t3
Time saved in second relearning.....	75.2 of t1	80.5 of t2	76.4 of t3
Time saved in third relearning.....	70.4 of t1	73.9 of t2	66.2 of t3

[1] Gordon: "Educational Psychology," page 169.

Of the three processes the best results are given by the second process, the saving of time in the relearnings being in all cases greater than in the other processes. Once more we notice the fact that the results between the first and the third processes are strikingly similar.

It is interesting also to note that it takes about the same time to relearn the vocabulary after forty-two days as it takes after two days. That it is not any greater is due to the fact that a relearning took place between the two tests, which compensates for the length of time. We have here a good illustration of the usefulness of frequent relearnings.

TABLE VI.—ACCURACY IN THE LEARNING AND RELEARNING FOR EACH PROCESS, IN PERCENTAGES OF THE PERFECT SCORE

	Silently	Aloud	Aloud (written recall)
Accuracy for first learning.....	75.3	81.9	82.5
Accuracy for first relearning....	63.1	73.6	65.1
Accuracy for second relearning.	64.0	73.3	64.9
Accuracy for third relearning...	46.3	51.0	41.5

Again we see results similar to those obtained with regard to time. The second process gives better results than the other two. Another point, interesting to note, is that the loss in the third recall is practically equal to the loss in the second recall. That is, after ten days, the results seem to be as good as after two days. It must not be forgotten, however, that while no relearning had taken place between the first and the second recall, a relearning had taken place between the second and third recall. Thus the relearning compensates in a certain measure for the length of time. The fact that in all cases, it does compensate so closely is worthy of notice.

CORRELATIONS

Do the students who do well in using one method do well using the others also? While the correlations do not give a definite answer they are, nevertheless suggestive.

(*a*) ACCURACY

Correlation between the accuracy of the first process and the second..... .67
Correlation between the accuracy of the first process and the third..... .56
Correlation between the accuracy of the second process and the third..... .55

(*b*) TIME

Correlation between the time of the first process and the second	.55
Correlation between the time of the first process and the third	.61
Correlation between the time of the second process and the third	.42

(*c*) TIME AND ACCURACY

Correlation between time and accuracy in the first process	−.23
Correlation between time and accuracy in the second process	−.25
Correlation between time and accuracy in the third process	−.39

The first two groups of correlations, while not exceptionally high, still show fairly well that there are not two groups, one doing well with one process and the other doing better with the other process, but that to a rather large extent the students who did well in one also did well in the others.

The correlation between time and accuracy are all negative, showing that the element of time has little to do with the accuracy. These findings are in agreement with the opinion of the psychologists. Norsworthy and Whitley in the work already mentioned say: "The maxim 'easy come, easy go' has been firmly fixed with respect to memory. Recent experiments with both children and adults as subjects prove conclusively that the quick learner is not the quick forgetter. Children who learn quickly retain more on the average than those who learn slowly, both as tested by immediate and permanent memory."

Gordon in his *Educational Psychology* giving the results on an experiment on memory says: "It seems that the ones who learned the greatest number of words retain the highest percentage at the last trial."

CONCLUSION

In the teaching of foreign languages, more and more emphasis is placed upon the necessity of teaching the student how to speak the language. This is is not an easy task, for pronunciation which requires a very accurate adjustment of muscles is a difficult task in the early stage of word learning. Furthermore, there is according to Handchin, "a special hindrance encountered in acquiring a second language, due to the extraordinary resistance of the native speech center." Watson explains this resistance in saying that "the habitual type of response robs the organism of its muscular flexibility."

To overcome this resistance much practice is required, for pronunciation is a matter of motor development, as well as auditory per-

ception and memory. Indeed, Palmer goes even farther when he says that "It is impossible to memorize speech material without articulating it in some form or other." Learning a language is then a matter of sensori-motor learning and the question of pronunciation is simply one form of the general problem of how voluntary motions are acquired. According to Freeman, "The golden rule of sensori-motor learning is much repetition." It follows clearly from the analysis of sensori-motor learning that progress can be gained only through a large amount of practice.

How are the students to get the necessary amount of practice? The time in the classroom is limited, and due to the number of students in the class, each one has little opportunity for individual practice. Furthermore, the amount of time devoted to the pronunciation proper is also very small. Indeed in the average classroom, the course is begun with a study of phonics so that the student may be given a clearer idea of the sounds, and that is about all that is done. This is far from being sufficient, for sounds are difficult not only in themselves, but according to the sounds with which they are associated, and it is only by practicing these combinations that the student will gain the necessary control of its vocal organs. Unfortunately, with the pressure put on the teacher to cover a certain program in so many months, it is difficult to find time for an intermediary stage between the pronunciation of the isolated sounds and their pronunciation in combinations. This very necessary stage may be supplied to some extent by urging the student to study aloud. However, could the objection be raised that such practice, excellent for pronunciation, might not prove to be the most economical way of memorizing foreign vocables? This experiment was undertaken in order to ascertain what process of study gave in reality the best results. Its findings are to the effect that studying aloud is the best method and that it has a double advantage, greater accuracy and more persistent retention. When these advantages are added to the undisputed increase in fluency in pronunciation it is clear that studying aloud is the most advantageous course to follow.

CORRELATION WITHOUT PLOTTING

JACOB S. ORLEANS

Research Associate, Educational Measurement Bureau
New York State Department of Education

The method of computing correlation coefficients described here is meant particularly for those who need to compute intercorrelations, or who on general principles wish to avoid the plotting necessary for the more commonly used methods. The use of this method has certain recognized limitations that are mentioned below. A number of short cuts are included in the procedure, but they are not indigenous to it. They may be applied to any work involving the computation of a mean or standard deviation of a grouped distribution.

The following notation will clarify the explanation that follows:

X and Y = measures of two variables
M_x and M_y = the respective means of these measures
$x = X - M_x$
$y = Y - M_y$
$\delta = x - y$ (for any one individual)
$D = X - Y$ (for any one individual)
M_d = the mean of the D measures
$d = D - M_d$
σ_x, σ_y and σ_d = the respective standard deviations of the X, Y, and D distributions

The usual Pearson formula for computing a coefficient of correlation may be transmuted in many different ways.[1] The following leads to the formula used here:

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2}\sqrt{\Sigma y^2}} \tag{1}$$

$$\delta = x - y$$

$$\delta^2 = x^2 - 2xy + y^2$$

$$\Sigma\delta^2 = \Sigma x^2 - 2\Sigma xy + \Sigma y^2$$

$$\Sigma xy = \frac{\Sigma x^2}{2} + \frac{\Sigma y^2}{2} - \frac{\Sigma\delta^2}{2}$$

$$\therefore r = \frac{\Sigma x^2 + \Sigma y^2 - \Sigma\delta^2}{2\sqrt{\Sigma x^2}\sqrt{\Sigma y^2}}$$

[1] Dr. P. M. Symonds of Teachers College, Columbia University, has worked out 52 such variations.

Dividing numerator and denominator by n gives

$$r = \frac{\delta_x}{2\delta_y} + \frac{\delta_y}{2\delta_x} - \frac{\frac{1}{n}\Sigma\delta^2}{2\delta_x\delta_y} \tag{2}$$

This is the familiar difference formula, which however is usable only when $M_x = M_y$. Otherwise the computation of the term $\Sigma\delta^2$ is cumbersome and time consuming. The following derivation avoids this term and substitutes for it the differences between the gross scores for the differences between their deviations from their respective means.

$$\begin{aligned}
\delta &= x - y = (X - M_x) - (Y - M_y) = (X - Y) - (M_x - M_y) \\
\delta^2 &= (X - Y)^2 + (M_x - M_y)^2 - 2(M_x - M_y)(X - Y) \\
\Sigma\delta^2 &= \Sigma D^2 + nM^2{}_x + nM^2{}_y - 2nM_xM_y - 2M_x\Sigma X + 2M_x\Sigma Y + \\
&\quad 2M_y\Sigma X - 2M_y\Sigma Y
\end{aligned}$$

but $\Sigma X = nM_x$ and $\Sigma Y = nM_y$

$$\begin{aligned}
\therefore \Sigma\delta^2 &= \Sigma D^2 + nM^2{}_x + nM^2{}_y - 2nM_xM_y - 2nM^2{}_x + 2nM_xM_y + \\
&\quad 2nM_xM_y - 2nM^2{}_y \\
&= \Sigma D^2 - nM^2{}_x + 2nM_xM_y - nM^2{}_y \\
&= \Sigma D^2 - n(M_x - M_y)^2
\end{aligned}$$

Substituting this value in (2)

$$r = \frac{\sigma_x}{2\sigma_y} + \frac{\sigma_y}{2\sigma_x} - \frac{\frac{\Sigma D^2}{n} - (M_x - M_y)^2}{2\sigma_x\sigma_y} \tag{3}$$

This formula is convenient if the values of D (the differences between the pairs of measures) are small, say, less than 20. But if many of them are large the value of the term $\frac{\Sigma d^2}{n}$ may be obtained more readily from a grouped distribution of the D measures by the following formula

$$\frac{\Sigma D^2}{n} = \sigma_d^2 + M_d^2$$

which is derived as follows

$$\sigma_d = \sqrt{\frac{\Sigma d^2}{n} - c_d^2} \tag{4}$$

This is the usual formula for the standard deviation of a distribution in which

σ_d = the standard deviation of the distribution of D's
$d = D - M_d$
M_d = the mean of the distribution of D's
c_d = the correction, the difference between M_d and the guessed mean

If the guessed mean is taken at 0, formula (4) becomes

$$\sigma_d = \sqrt{\frac{\Sigma D^2}{n} - M_d^2}$$

or

$$\sigma_d^2 = \frac{\Sigma D^2}{n} - M_d^2$$

or

$$\frac{\Sigma D^2}{n} = \sigma_d^2 + M_d^2$$

Substituting this value for $\frac{\Sigma D^2}{n}$ in (3)

$$r = \frac{\sigma_x}{2\sigma_y} + \frac{\sigma_y}{2\sigma_x} - \frac{\sigma_d^2 + M_d^2 = (M_x - M_y)^2}{2\sigma_x\sigma_y} \qquad (5)$$

This formula is much more formidable looking than the original Pearson formula. However, the processes involved are simpler and under certain circumstances, *e.g.*, computing intercorrelations, a large part of the labor is saved. The actual work consists of making three distributions and computing their means and standard deviations.

The steps in the work are

1. Make the X distribution
2. Make the Y distribution
3. Make the D distribution
4. Compute M_x, M_y, M_d
5. Compute σ_x, σ_y, σ_d
6. Substitute these values in the formula

If, as stated above, the values of D are small it is not necessary to make a grouped distribution and compute the mean and sigma. The value of $\frac{\Sigma D^2}{n}$ can be computed directly by squaring each difference and summating, and substituting in the shorter formula (3).

An illustration of the use of the formula follows. It is advisable, if feasible, to obtain the data so that the scores in one trait are arranged in order from lowest to highest. Besides facilitating the computation this allows for an estimate of the size of the correlation.

The steps in the computation and substitution in the formula are given in the usual job analysis form in the chart.

Correlation between and r = .624 Date

X	Distribution	+f	−f	x	+fx	−fx	fx²	Y	Distribution	+f	−f	y	+fy	−fy	fy²
				12								12			
				11								11			
				10								10			
44	II	2		9	18		162					9			
42	I	1		8	8		64					8			
40	I	1	1	7			98	70	I	1	2	7		7	147
38	IIII	4	−	6	24		144	65	I	1	2	6		6	108
36	III	4	3	5	5		175	60	II	2	3	5		5	125
34	IIII II	7	6	4	4		208	55	II	2	6	4		16	128
32	IIII I	6	9	3		9	135	50	IIII	5	10	3		15	135
30	IIII I	6	14	2		16	80	45	IIII II	7	18	2		22	100
28	IIII III	8	15	1		7	28	40	IIII IIII	16	18	1		8	18
26	IIII IIII III	13		0	59	32		35	IIII IIII III	13		0	.79		
24	IIII IIII IIII	15		-1				30	IIII IIII IIII	18		-1			
22	IIII IIII	14		-2				25	IIII IIII IIII	18		-2			
20	IIII IIII	9		-3				20	IIII IIII	16		-3			
18	IIII I	6		-4				15	IIII I	6		-4			
16	III	3		-5				10	III	3		-5			
14		−		-6				5	II	2		-6			
12	I	1		-7				0	II	2		-7			
				-8								-8			
				-9								-9			
				-10								-10			
				-11								-11			
				-12								-12			
G_x =	27.0	100			27	1089		G_y =	37.5	100			-79	761	
c_x =	2	A			B	C		c_y =	5	A			D	E	

D	Distribution	+f	−f	d	+fd	−fd	fd²
				12			
				11			
				10			
				9			
				8			
				7			
28	II	2	2	6			144
24	III	3	4	5		5	125
20	IIII	7	5	4	8		192
16	IIII	5	4	3	3		81
12	IIII II	8	4	2	8		48
8	IIII IIII I	21	19	1	2		40
4	IIII IIII I	16	0	0	21	5	
0	IIII IIII	19	-1				
-4	IIII	4	-2				
-8	IIII	4	-3				
-12	IIII	5	-4				
-16	IIII	4	-5				
-20	II	2	-6				
			-7				
			-8				
			-9				
			-10				
			-11				
			-12				
G_d =	5.5	100			+16	680	
c_d =	4	A			F	H	

B+A	.27	I
I·c_x	.54	J
G_x·J	27.54	M_x
D÷A	−.79	K
K·c_y	−3.95	L
G_y+L	33.55	M_y
M_x−M_y	6.01	m

C+A	10.89	N
I^2	.0729	O
N−O	10.8171	P
$\sqrt{P}$	3.29	Q
Q·c_x	6.58	R

E÷A	7.61	S
K^2	.6241	T
S−T	6.9859	U
$\sqrt{U}$	2.64	V
V·c_y	13.21	W

F÷A	−.16	X
X·c_d	−.64	Y
G_d·Y	6.14	M_d
H÷A	680	Z
X^2	.0256	a
Z−a	6.7744	b
b·c_d^2	108.39	e
M_d^2	37.70	f

e+f	146.09	g
m^2	36.12	h
g−h	109.97	i
±RW	173.84	j
±j	.634	k
l^2	1.250	l
l−k	.616	n

*l is obtained from Table I

CHART 1.

X	Y	X	Y	X	Y	X	Y	X	Y
13	16	22	40	24	47	27	15	33	34
16	5	22	25	25	37	28	28	34	34
16	20	22	31	25	23	28	28	34	62
17	22	22	12	25	40	28	44	34	29
18	25	22	31	25	33	28	26	34	35
18	25	23	23	25	37	28	39	35	58
19	29	23	25	25	37	29	43	35	49
19	31	23	31	25	42	29	50	35	44
19	17	23	31	26	26	29	37	36	36
19	2	23	8	26	36	30	25	36	37
20	14	23	29	26	24	30	18	36	46
20	27	23	32	26	28	30	36	37	41
20	21	23	43	26	34	31	48	38	54
20	30	24	24	27	34	31	56	38	26
20	3	24	17	27	34	31	35	39	24
21	21	24	33	27	52	32	39	39	45
21	23	24	33	27	33	32	43	41	50
21	27	24	11	27	47	32	62	43	66
21	32	24	26	27	36	32	54	44	70
22	32	24	28	27	41	33	19	45	47

TABLE I.[1]—VALUES OF $\frac{\sigma_x}{2\sigma_y} + \frac{\sigma_y}{2\sigma_x}$ FOR GIVEN VALUES OF $\frac{\sigma_x}{\sigma_y}$ (ONLY PART OF THE TABLE IS SHOWN)

$\frac{\sigma_x}{\sigma_y}$	$\frac{\sigma_x}{2\sigma_y} + \frac{\sigma_y}{2\sigma_x}$	$\frac{\sigma_x}{\sigma_y}$	$\frac{\sigma_y}{2\sigma_y} + \frac{\sigma_y}{2\sigma_x}$
1.0	1.000	2.0	1.250
1.1	1.005	2.1	1.288
1.2	1.017	2.2	1.327
1.3	1.035	2.3	1.367
1.4	1.057	2.4	1.408
1.5	1.083	2.5	1.450
1.6	1.113	2.6	1.493
1.7	1.144	2.7	1.535
1.8	1.178	2.8	1.579
1.9	1.213	2.9	1.622

This table reads as follows: if the ratio of σ_x and σ_y $\left(\text{either } \frac{\sigma_x}{\sigma_y} \text{ or } \frac{\sigma_y}{\sigma_x}\right)$ is 1.0 then $\frac{\sigma_x}{2\sigma_y} + \frac{\sigma_y}{2\sigma_x}$ equals 1.000; if the ratio of the sigmas is 1.1 then $\frac{\sigma_x}{2\sigma_y} + \frac{\sigma_y}{2\sigma_x}$ equal 1.005, etc.

In the present illustration, divide 13.21 by 6.58 (always divide the larger sigma by the smaller). The ratio is 2.001. The value of $\frac{\sigma_x}{2\sigma_y} + \frac{\sigma_y}{2\sigma_x}$ corresponding to this is 1.250, which is the value of l in the job analysis chart.

[1] I am indebted to Dr. A. S. Otis of the World Book Co. for suggesting the present form of this table in place of the more cumbersome one originally used.

The determination of the class intervals and guessed means is a matter of practice. If the class interval is not properly chosen at first, a little time is lost. If the guessed mean is taken too far from the actual mean, the computation will involve larger numbers than it is necessary to deal with.

The notation on the chart has the following significance:

G_x, G_y, G_d[1] = Guessed means of X, Y, and D distributions respectively
c_x, c_y, c_d = class intervals of the X, Y, and D distributions respectively
$A = n$ (number of cases)
$B = \Sigma x$ (x here meaning deviation from G_x)
$C = \Sigma x^2$
$D = \Sigma y$ (y here meaning deviation from G_y)
$E = \Sigma y^2$
$F = \Sigma d$ (d here meaning deviation from G_d)
$H = \Sigma d^2$

The other letters represent the steps in the computation necessary to find the three means and the three sigmas and to substitute these in the formula. The value of l (in the last column of the computation) is obtained from Table I.

Since in the present illustration the differences are small, it is just as simple to obtain $\frac{\Sigma D^2}{n}$ directly by squaring the differences and summating. This may be done by punching into an adding machine the square of each difference as it is read off, or by making an ungrouped distribution (disregarding signs since the differences are to be squared) and multiplying the square of each difference by its frequency. The work would appear as follows:

---

[1] The guessed mean of the D distribution must be taken at the midpoint of the interval assuming the scores to run from $-.5$ to .5, .5 to 1.5, 1.5 to 2.5, etc. instead of from 0 to 1, 1 to 2, 2 to 3, etc.

D	f	fD^2
30	1	900
29		
28	1	784
27		
26	1	676
25	2	1,250
24		
23	3	1,587
22	1	484
21	1	441
20	2	800
19		
18	1	324
17	4	1,156
16	2	512
15	3	675
14	4	784
13	1	169
12	7	1,008
11	4	484
10	6	600
9	8	648
8	5	320
7	7	343
6	6	216
5	3	75
4	4	64
3	2	18
2	9	36
1	4	4
0	8	
	100	14,538

Therefore $\frac{\Sigma D^2}{n}$ equals 143.58. This is 2.51 less than the value obtained above and would raise the correlation from .616 to .632, a negligible difference.

This direct method of obtaining the value of $\frac{\Sigma D^2}{n}$ by an ungrouped distribution is shorter if, as indicated, the differences are small. Much shorter is the more direct addition of the squares of the differences on an adding machine, which takes hardly more time than is necessary to write the distribution of the differences. As a matter of fact, if the differences are large a slight error will make an appreciably large change in the square and the size of the correlation may be seriously affected. Likewise if as a result of error due to coarseness of grouping, the difference between the means is even 2 points too large or too small, and the difference between the means is a large quantity, the correlation may be affected to the extent of .10 or .20. Thus if the difference between the means is computed as 92 instead of 90, as the result of coarse grouping, the numerator of the last fraction of the formula will be 364 too small. If the denominator is 30, the correlation will be .12 too large. In other words it is wiser not to use this correlation procedure if the differences between the scores or their means are large. The formula in such instances is too sensitive.

In working the above sample correlation several short cuts are employed that may be used in the computation of a mean or standard deviation from any grouped distribution. The customary procedure in such work is to multiply each deviation from the guessed mean by its frequency, add the positive products, add the negative products, and take the difference between these two sums; and to multiply the square of each deviation by its frequency and add the sum of these squares. In the present illustration, the frequencies for the negative deviations are copied beside the positive frequencies of the deviations of equal magnitude. The deviation is then multiplied by the differ-

ence between the frequencies, the product being written in the positive column if the positive frequency is the larger and in the negative column if the negative frequency is the larger; and the square of the deviation is multiplied by the sum of the frequencies, the product being written in the fx^2 column. It is obvious then, particularly with a large number of cases, there is a large saving of labor and time in that fewer operations have to be performed and smaller and fewer numbers have to be dealt with. It often happens that the frequencies for the same sized deviations are equal, or that the sum of the frequencies is a multiple of 10, or is some other convenient number.

The disadvantages of using the correlation procedure described here may be listed as follows: (1) There is no plot or "scattergram" and therefore it is not possible to tell whether the correlation is linear without further investigation, nor to estimate easily the probable size of the correlation. (2) If the differences vary a great deal and many of them are large, the computation of $\frac{\Sigma D^2}{n}$ becomes cumbersome. (3) The formula is too sensitive for large values of D, and for large values of the difference between the means of the X and Y distributions.

The advantages of using this method are: (1) It is not necessary to make a plot, thus avoiding a tedious, eye-straining task, and also avoiding the computation of Σxy or adding diagonals. (2) The computation involves only standard deviations and means, items familiar to those whose work involves statistical treatment. (3) If either of the measures is to be used in another correlation, it is not necessary to compute its mean and sigma again. If the intercorrelations for, say, 6 variables are to be computed, it is necessary to make six distributions, one for each variable, and 15 distributions of differences between each pair of variables. By a plotting method the distribution for each variable would be made five times.

The use of the above described correlation method will, then, be found advantageous for data having a fair number of cases, say, at least 100, and in which the differences between the scores are small. It is particularly useful in working intercorrelations.

PLAY ACTIVITY AND SCHOOL PROGRESS

HARVEY C. LEHMAN AND PAUL A. WITTY

The University of Kansas

Educational writers have emphasized repeatedly the fact that normal play life is essential if the child is to develop to his maximum. Within recent years there has been much discussion regarding the advisability of separating the gifted child from companions of his own age in order that he may progress in his school work at a rate consonant with his mentality. Prominent writers have taken opposite stands on this question. Holmes has argued that the gifted child should not be allowed to "run the dangers of a forced pace during the earlier years."

"Something should be said for normality. Health, companionship, and happy participation in the activities of his companions are considerations which should all be taken into account in dealing with every individual case. Education is a means whereby the individual may have full development among his fellows and for the common good. No short-sighted view of what individual development means should lead us to separate a bright child from the companions with whom he can be happiest and from whom he can learn most through common work and play.

". . . Nature has a program in the development of children of which we must also take account, and it may be far better to curtail or telescope the higher stages of education, which come after natural development is more nearly completed, than to run the dangers of a forced pace during the earlier years.

". . . Before we assume that they (gifted children) ought to be encouraged to complete their work in the grades and in the high school in less than the usual time, we ought at least to experiment with the plan of allowing them, instead, to use the time they have on school routine in freer, happier and more rewarding ways."[1]

Klapper too has insisted that the gifted child should not be separated from his companions. He states that maladjustments result when a child is permitted to progress through the grades too rapidly.

"The junior high school graduate of eleven is a maladjusted child. Child life is more than mere school study. The eleven-year-old graduate cannot play with children eleven years old, nor can they play safely with those fourteen-year-old children who are their mental equals. The high-school senior of fourteen finds few friends among his class-mates of seventeen. He is decidedly a maladjusted adolescent and the school has unwittingly become a part in producing his problem . . . Would it not be wiser to keep the superior child in school longer and to enrich its education by the addition of a variety of manual and trade experiences by music, by physical activities, by club life, by visits to

museums, and by extensive reading of fiction, of current events, and of biography? Where is the enriched curriculum that is so frequently promised?"[2]

The preceding assertions are interesting but their subjective nature renders them of questionable validity.

Freeman, on the other hand, has emphasized that acceleration as a means of adjustment has advantages not generally recognized or accredited to it; that grade skipping as a regular practice is not as detrimental as we have been inclined to think it; its results being dependent upon the ability of the pupil. He states:

"We have in the past emphasized the distinction between acceleration and enrichment. In so doing, we have made a false distinction. The real distinction is between the adjustment which merely aims at saving time and the adjustment which aims at securing for the gifted an opportunity to do work at a higher intellectual level. We have further assumed that enrichment implies keeping the pupil engaged in the work in which the pupil of average classification of the same age is engaged. This assumption is, I believe, incorrect. Acceleration actually provides enrichment. The work of the advanced grades is intellectually superior because the method which is pursued and the content are superior to those of the lower grades. From the point of view of intellectual adjustment, then, acceleration accomplishes both the saving of time and the enrichment of the instruction. The difficulties with this mode of enrichment are not of an intellectual nature but of a social nature. These difficulties may be met by proper forms of organization and they are being progressively diminished by the very increase in the frequency of acceleration itself."[3]

The above quotations give evidence of conflicting opinions regarding the effect of acceleration upon the child's adjustment. Data in this regard are scant. The present writers have attempted to obtain salient data in respect to the effect of acceleration upon the play life of the child.

The data herein presented disclose (1) the relationship between school progress and the number of play activities participated in by representative school children of Kansas City, Missouri, and (2) the extent to which retarded and accelerated pupils engage in solitary play activities.

Method

The Lehman Play Quiz was administered to more than 6000 children in Grades III–XII in the public schools of Kansas City, Missouri. The Play Quiz was devised to cover these specific grades. The children are asked to indicate among a comprehensive and catholic list of 200 activities *only those* in which they have engaged *during the*

preceding week. The children are later asked to designate the activities in which they participated *alone.* These activities in which the children engage without companionship will be designated solitary activities when alluded to in the discussion following.

Progress quotients were obtained by dividing the mean CA in months of Kansas City pupils of a given grade by the CA in months of each child in that grade.

For each child the total number of play activities engaged in during the preceding week was ascertained. The number participated in in company with one or more additional children was next determined. The per cent of the total activities that these social activities represented was designated the index of social participation. Thus an index of social participation of 80 indicates that 80 per cent of the activities engaged in by the child were ones in which one or more other children also took part.

TABLE I.—PLAY DATA FOR 6886 CHILDREN

CA	Frequencies	Mean number of activities engaged in	Mean index of social participation
7½	84	44.26	62.01
8½	468	40.56	63.25
9½	935	42.37	61.70
10½	981	37.67	60.58
11½	748	36.86	58.12
12½	903	34.01	55.69
13½	946	31.52	55.65
14½	848	28.58	52.92
15½	573	27.45	52.28
16½	288	25.91	50.56
17½	82	24.93	52.04
18½	25	25.50	52.32
19½	5	25.50	57.50
Total.....	6886		

The data were assembled for 6886 children. These were grouped according to increasing CA in six months intervals. Table I presents the following data for the children in each age interval: (1) Frequencies (2) Mean number of activities engaged in and (3) Mean indices of social participation. The first significant finding was the fairly con-

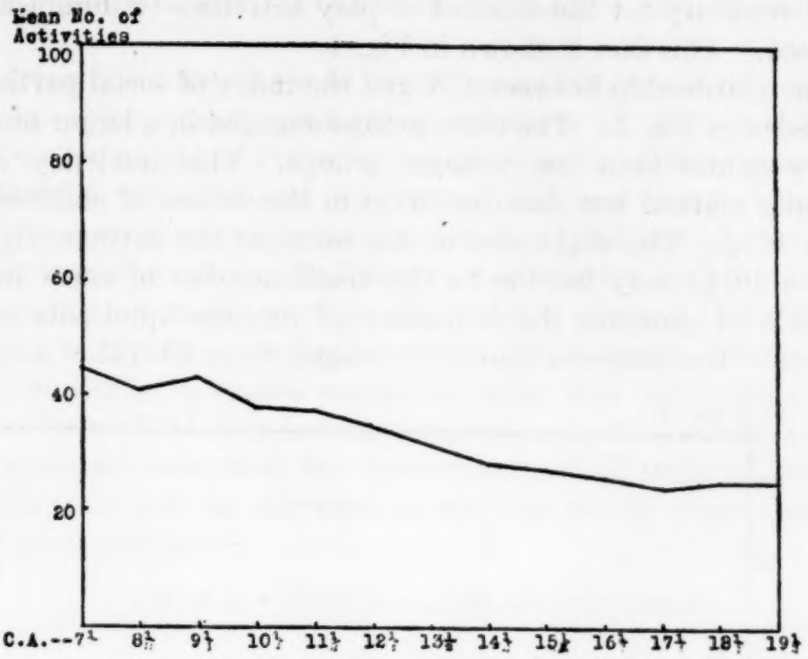


FIG. 1.—Relationship between versatility of play interests and CA.

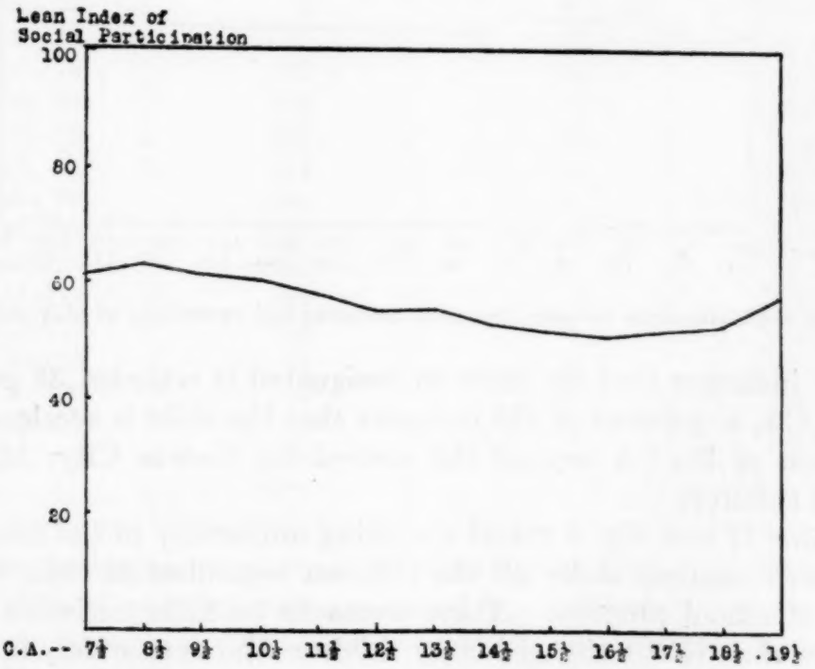


FIG. 2.—Relationship between CA and the index of social participation.

sistent tendency for the number of play activities to diminish as CA increased. This fact is shown in Fig. 1.

The relationship between CA and the index of social participation is revealed in Fig. 2. The older groups engaged in a larger number of solitary games than the younger groups. This tendency was not especially marked but was consistent in the indices of children of CA 8½ to 16½. The slight rise of the curve at the extreme right (CA 17½ to 19½) may be due to the small number of cases involved.

Table II presents the frequency of progress quotients for 6374 children. The progress quotients ranged from 65–135. A quotient

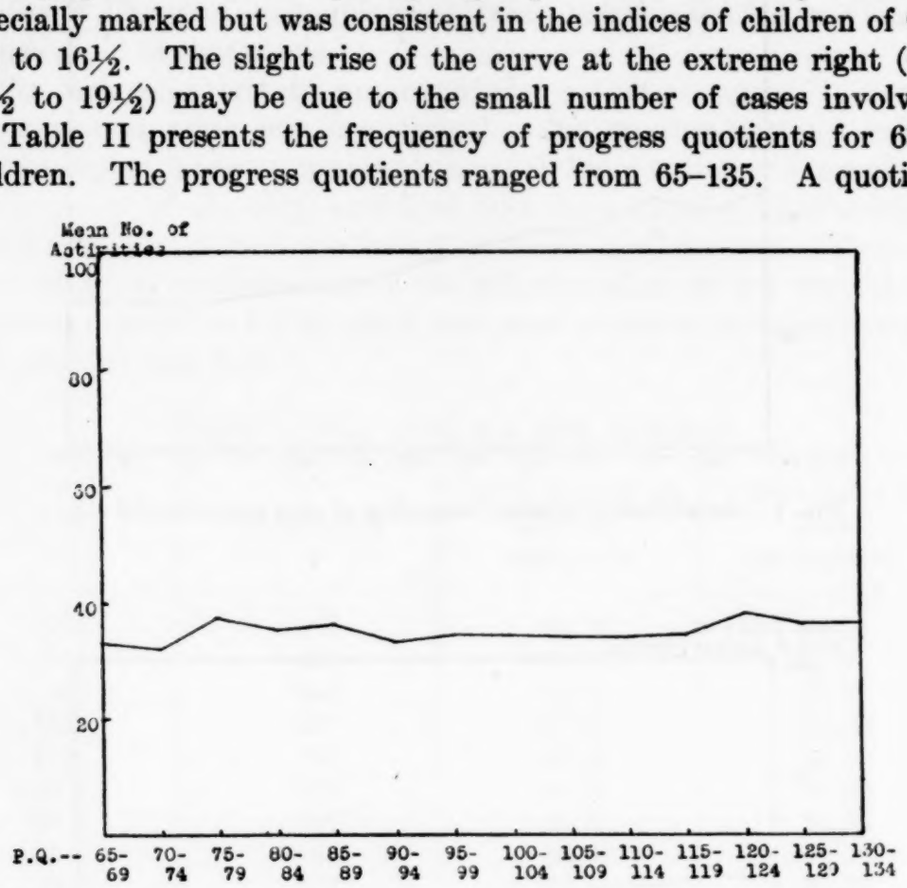


FIG. 3.—Relationship between progress quotients and versatility of play interests.

of 65 indicates that the child so designated is retarded 35 per cent of his CA, a quotient of 135 indicates that the child is accelerated 35 per cent of his CA beyond the normal for Kansas City, Missouri, school children.

Table II and Fig. 3 reveal a striking uniformity in the number of activities engaged in by all the children regardless of their varying rates of school progress. There seems to be little variation in the *number* of activities engaged in by children who vary widely in rate of progress in school. Between the intervals represented by progress quotients 80–120, there was found striking regularity. The extremes

(progress quotients 120–135) also adhered rather closely to the general trend.

There are too few cases to permit generalization in regard to the extremes. The meagre evidence in reference to them corroborates, however, the results of the more complete body of data; namely, the educationally retarded and accelerated children engage in *approximately the same number of play activities as do children who are making normal progress in their school work.* Acceleration does not exact its toll in reduction in the number of activities engaged in by the child after this adjustment has been made. Nor does the child who progresses normally in school engage in more play activities than the retarded one. The data herein presented seem to show the falsity of the traditional view that the accelerated school child demonstrates a conspicuous lack in reference to the number of plays and games in which he participates.

TABLE II.—PLAY DATA FOR 6374 CHILDREN

Progress quotients	Frequencies	Mean number of activities engaged in	Mean index of social participation
65– 69	8	33.20	75.31
70– 74	31	32.19	73.95
75– 79	75	37.81	68.86
80– 84	195	35.36	62.64
85– 89	363	36.07	61.82
90– 94	663	33.49	59.11
95– 99	999	34.87	58.67
100–104	1427	34.50	56.03
105–109	1482	34.38	57.53
110–114	868	34.10	55.89
115–119	182	34.66	54.28
120–124	62	38.15	57.82
125–129	12	36.43	57.50
130–134	7	36.67	61.08
Total.....	6374		

The foregoing data have revealed no appreciable difference in versatility of play interests exhibited by children of widely varying progress quotients. The next important consideration is concerned with the types of play in which these children participated. Particularly salient is the analysis of play in reference to the extent to

which social participation is permitted. This relationship is made evident in the indices of social participation.

Table II and Fig. 4 present the relationship between indices of social participation and progress quotients. In contrast to the slight differences found in the social characteristic of play for children of increasing CA, there existed a rather marked tendency for pupils having very low progress quotients to turn to social play activities. It will be noted from Table II that the index of social participation decreases as the progress quotient increases. This tendency was

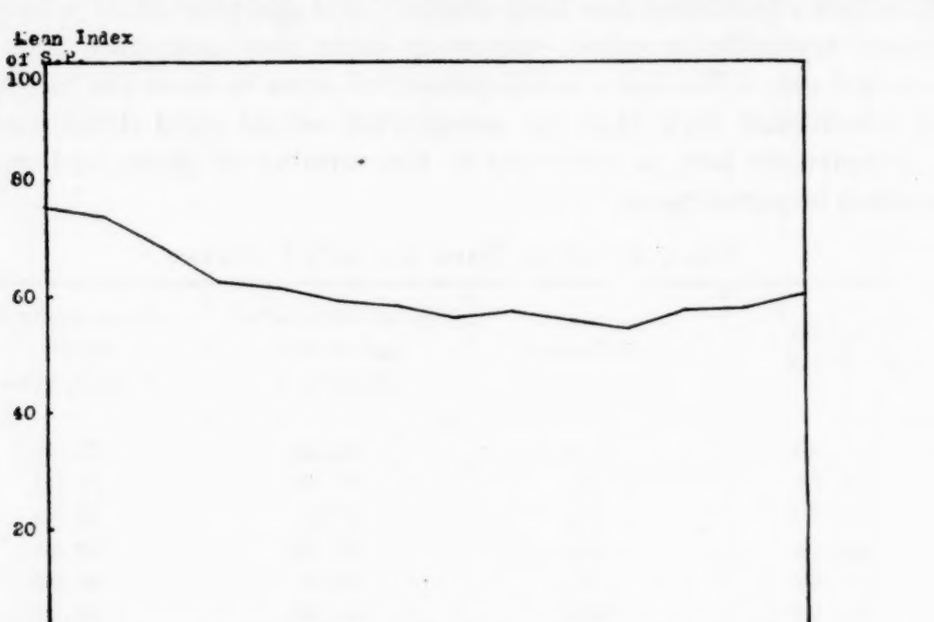


FIG. 4.—Relationship between the index of social participation and progress quotients.

particularly noticeable with children of progress quotients 65–100. Fifty-six per cent of the activities engaged in by children of progress quotient 100, and 75 per cent of the activities participated in by those of progress quotient 65 were social ones. No appreciable difference was discernible in the type of play in which children of progress quotients 100–135 take part.

Robinson points out that it is unsafe to assume a causal relationship just because a correlation is found. Referring to the negative correlation that is always found between smoking and school marks, he expresses the following:

> Correlations between school grades and smoking students prove nothing. Groups of smokers always make poorer marks than groups of non-smokers. The same lack of social inhibition that causes the boy to smoke may also cause him to neglect his studies. The same social environment that causes the older boy to take up smoking may also offer a maximum number of distractions from purely scholastic activities.[4]

It is difficult to account for the fact that the pupils of lowest progress quotients are found to be most social in their play. It is possible that the very environment that provides choice companionship also offers a maximum number of distractions from purely scholastic activities.

Summary

Play data were assembled for 6886 children grouped according to their varying rates of progress in school. The number of play activities engaged in and the mean indices of social participation were obtained for the various groups.

There was no appreciable difference revealed in the study made by the present writers in the number or diversity of play activities engaged in by children of widely varying progress quotients.

There was a considerable variation among children of varying progress quotients with respect to the *type* of play activity engaged in.

Pedagogically retarded children participated in a considerably larger number of *social* play activities than did children who had progressed normally in school. It seems reasonable to assume that the type of activity in which the child spontaneously and voluntarily participates in is a reflection of his felt need. The pedagogically retarded child then demonstrates a conspicuous need for and active interest in plays and games which provide social contacts. It is of course possible that the pedagogically retarded child's felt need is an inadequate criterion. The question is one that cannot be answered until further studies have been made. There is need for careful case studies of the personality development of the pedagogically retarded child. Such studies will need to follow the child's development, and his success as an adult in the game of life will need also to be taken into account. It is possible that the retarded child's social development compensates somewhat for his pedagogical retardation; the validity of this assumption will need to be substantiated by subsequent investigation.

Pedagogically accelerated children showed little variation from children who had progressed normally in school in reference to the

social quality of their play behavior. They did not engage in a smaller number of solitary plays and games. On the contrary, there was a tendency for pedagogically accelerated children to have engaged in a slightly larger number of social play activities than children who had progressed normally. The difference between the two groups, however, was not marked in this regard. The significant fact is that acceleration did not effect a reduction either in the *number of different play activities engaged in or the extent to which the child participated with others in recreational activities.*

References

1. Holmes, H. W.: Intelligence Tests and Individual Progress in School Work. *Twenty-second Yearbook of the National Society for the Study of Education.* Bloomington, Illinois. The Public School Publishing Company. Part I, pp. IX + 275 (pp. 121–122).
2. Klapper, Paul: The Experimental Study of Education with Special Reference to the Elementary School. *Journal of Educational Research,* Vol. XII, September, 1925, pp. 126–127.
3. Freeman, Frank N.: The Treatment of the Gifted Child in the Light of Scientific Evidence. *The Elementary School Journal,* May, 1924, pp. 652–661.
4. Robinson, Edward S. and Robinson, Florence Richardson: "Readings in General Psychology." Chicago. The University of Chicago Press, 1923, pp. XVI + 674 (p. 655f).

THE RELATIVE VARIABILITY OF BOYS AND GIRLS*

A. LEON WINSOR, M.A.

From the Department of Psychology, Stanford University

In *Genetic Studies of Genius*[38] boys are found far to outnumber girls in the higher IQ ranges. This agrees with the supposedly well-established fact of the greater variability of males. In Terman's study a school population of one quarter of a million was canvassed and the highest 1 per cent was selected for study. The various factors used as a basis for selection were intelligence tests, achievement tests, age-grade status, and teachers ratings. Special care was, of course, taken to eliminate sex preference in these selections. In spite of this care the sample finally taken exhibited an unusual ratio of males to females. In the three groups studied the following distribution obtains:

	Boys	Girls	Ratio
Group I (the Main Group)	352	291	121 to 100†
Group II (the Outside Group)	197	159	123 to 100
Group III (the High School Group)	257	120	212 to 100

Concerning the preponderance of males Terman and his collaborators have the following to say:

"The true cause of the sex ratio found can not be determined from our data. It may be either variability or the differential death rate of embryos. Both of these factors may be involved and possibly others. Biased selection due to the method of nomination and testing is probably not responsible—except for the non-excess of boys in the highest IQ ranges the facts we have presented are in harmony with the hypothesis that exceptionally superior intelligence occurs with greater frequency among boys than among girls." [38, page 54]

It is our purpose to examine this hypothesis in the light of data now available. The relative variability of males and females has been a subject for many discussions in the past and recognized authorities in the field of psychology have participated on each side of the question. Often the assumptions made by the participants have been based upon inadequate data. Differences in variability between five men and five women have been cited as evidence of greater male variability. In only a few cases have figures of an impressive character been presented. In 1914 Hollingworth[16] gave an account of the

* Communicated by Professor Walter R. Miles.

† The ratio of males to females in living births of whites in the U. S. is 105.9 to 100. J. B. Nickols: Memoirs of American Anthropological Association. Vol. I, 1907, pp. 249–300.

history of the discussion and reviewed the studies that had been made previously. She summarizes her findings in part as follows:

1. "The greater variability of males in anatomical traits is not established but is debated by authorities of equal competence.

2. "But even if it were established it would only suggest, not prove, that men are more variable in mental traits also. The empirical data at present available on this point are inadequate and contradictory, and if they point either way actually indicate greater female variability.

3. "But even if it were established that there actually is greater male variability in mental traits, it would only suggest, not prove that there is greater inherent variability. For (*a*) the opportunity and exercise of the sexes have been dissimilar and unequal; (*b*) intellectual variability has had survival value for men, but for women it has had little or none—this by virtue of the different parts played by the sexes in the perpetutation of the species."

Some extensive studies in sex differences have been made since Hollingworth's review. Unfortunately the phase of the subject with which we are concerned has either been neglected or dismissed without serious consideration. The data, however, are of such a nature that the contents can be readily applied for comparative purposes. In the early development of the hypothesis Havelock Ellis[8] based his claim of greater male variability upon anatomical and pathological phenomena. This aspect of the subject will therefore be considered first.

To answer the question, "Are male human beings inherently more variable in anatomical traits than female human beings?" Montague and Hollingworth[18] examined the anatomical measurements of 2000 infants born in the obstetrical wards at the New York Infirmary for Women and Children. The measurements considered were weight, length, circumference of shoulders, and seven different cranial measurements. They were made immediately after birth before the child had ingested food. The examination included 1000 consecutive males and 1000 consecutive females. The assembled facts for all of the individual measurements showed:

1. Male infants are without exception, slightly larger than female infants in all anatomical measurements.

2. "If we take the gross average deviation as the measure of comparative variability, we find that in six cases (measurements) the males are more variable and in four cases the females are more variable.

3. "If we take the Pearson Coefficient as the measure of variability we find that in one case the males are more variable; in three cases the

females are more variable; and that in six cases there is no difference in variability when the PE of the coefficient is computed.

4. "In all cases the differences in variability are very slight. In only two cases does the percentile variation differ in the first decimal place."

The conclusion of Montague and Hollingworth is as follows: "Our results, showing that there is no inherent difference in anatomical variability, suggest that there will be found to be no inherent difference in mental variability, but they do not prove that such is the case. The following table is taken from their data and is inserted here for reference later in the discussion.

TABLE I.[1]—COMPARATIVE VARIABILITY IN BODY LENGTH AT BIRTH FOR MALE AND FEMALE BABIES

Sex	Number	Average	Average deviation	Pearson's coefficient of variation	PE of Pearson's coefficient
Males	1000	50.51	2.349	.0465	.0006
Females	1000	49.90	2.275	.0456	.0006

[1] Copied from Montague and Hollingworth. See their Table D, p. 354. (The measurements are given in centimeters.)

Anatomical factors were considered also by Henman and Livingston[14] in their study of comparative variability at different ages. They used the data of Burk and others to compute the variability of 88,449 individuals measured for height. Sex differences for different ages are as shown in Table II.

TABLE II.—COMPARATIVE VARIABILITY AT DIFFERENT AGES

Age	Boys, height	Girls, height	Variability difference in favor of boys	Variability difference in favor of girls
8.5	47.1	46.6	....	.003
9.5	49.7	48.7		
10.5	51.8	51.1	....	.002
11.5	53.5	53.1	....	.004
12.5	55	56.7	....	.005
13.5	56.9	57.8	.004	
14.5	59.3	59.8	.012	
15.5	62.2	60.9	.017	
16.5	64.3	61.7	.012	

Pearsons coefficient* of variability was the measure used here. Again the differences in variability are negligible. At those ages where the boys' average is a little the higher, the girls are slightly more variable. On the other hand when at adolescence the girls become somewhat taller than the boys, the latter show a slightly greater variability. This difference is probably due to the formula *used*, and is not significant. In these two studies of variability in anatomical traits where relatively absolute measurements are possible, there seems to be no foundation for the hypothesis in question. Slight differences have been found but they vary with the feature measured and the formula used.

In the field of mental testing extensive experiments and measurements are now available for analysis. The mentality survey of the elementary schools of Utah directed by George S. Snoddy affords data of especial interest here, although this phase of the subject was not considered in the report of the director. A total of 8088 pupils between the ages of 8 and 12 were given a modification of the Army Beta group test. The subjects represented the total population for these ages in 60 widely scattered schools. Each of the seven distinct tests for any given age yielded results which might be used here, but in the interest of brevity we shall limit our discussion to two of the more reliable tests, numbers 1 and 5. The responses of the 10-year-olds will be given.

Test 2 of the Army Beta, it will be recalled, involved the counting of 16 variously arranged piles of blocks. These piles were presented in the order of difficulty and the problem was to estimate the number of

TABLE III.—SCORES SHOWING THE COMPARATIVE VARIABILITY OF 10-YEAR-OLDS FOR THE ARMY BETA TEST 2

Sex	Age	Number[1]	Mean scores	Standard deviation	Coefficient of variability (P)
Boys	10	1161	13.12	4.80	.365
Girls	10	1160	10.06	4.08	.405

[1] The similarity in population in the two groups was a chance finding.

blocks in as many of the piles as possible in a given time. Two points were allowed for each pile correctly estimated. The tests were given under uniform conditions by carefully trained assistants. The scores for this performance were as indicated in Table III.

* Pearsons coefficient (CV) $= \frac{\text{The measure of variability SD or AD}}{\text{The central tendency mean or median}}$.

Test 5 consisted of comparing digits arranged in paired columns, and indicating on a black line whether or not the digits were identical. There were 50 such pairs arranged in accordance with their seeming difficulty. One point was allowed for each correct response. The score was the number right minus the number wrong. For the same group of 10-year-olds reported above, the following results were obtained. Only one individual (a girl) in the entire 10-year-old population made a perfect score.

TABLE IV.—SCORES SHOWING THE COMPARATIVE VARIABILITY OF 10-YEAR-OLDS FOR THE ARMY BETA TEST 5

Sex	Age	Number	Mean scores	Standard deviation	Coefficient variability (P)
Boys	10	1161	24.42	8.15	.333
Girls	10	1160	27.72	8.70	.317

The above scores were distributed as shown in Tables V.

TABLE V.—DISTRIBUTION OF THE SCORES SHOWN IN TABLE IV

Scores	Number of boys	Number of girls
45 or above	4	24
40 to 44	22	64
35 to 39	69	128
30 to 34	160	222
25 to 29	243	298
20 to 24	286	221
15 to 19	244	120
10 to 14	72	56
5 to 9	43	10
0 to 4	18	18

A glance at the means shows that in Test 2 the girls are 80 per cent as efficient as the boys while in Test 5 the boys are but 88 per cent as efficient as the girls. The Standard Deviation which was used here as the measure of gross variation shows the boys to be somewhat more variable than the girls in Test 2, while the same measure shows the girls to be more variable in Test 5. Applying Pearson coefficient of variation the situation is just reversed; the girls are shown to be more variable in Test 2 and the boys in Test 5. Although the differences by either measure are not large, attention must be paid to the

fact that the girls seem to be more variable in one performance and the boys in another. This same phenomenon, it will be recalled was shown when anatomical traits were under consideration. Furthermore the *measure of variability* used can be seen to play a determining rôle in the results obtained. That the variability as measured by the Pearson formula will be largely influenced by the differences in central tendency is obvious. A careful recheck of much of the data presented in evidence of greater male variability, however, shows not merely that this formula has been employed but that in the measures of performance used the girls have shown a higher average score than the boys. Thus this hypothesis might have developed because of the peculiarity of the formula frequently used, and the fact that in many of the tests used for measuring intelligence, memory factors and other elements now known to favor female performance have predominated. Touton (1924) in an article entitled, "Sex Differences in Geometric Abilities"[45] describes a critical study of 2800 New York High school pupils. The tests given were a part of the June Regents Examination List and the papers selected for study were random samplings of approximately 500 solutions written by each sex for each exercise. Since a large number made perfect scores in some of these tests, only those exercises will be considered whose per cent of perfect scores is not excessive. Problems 9 and 11 had few perfect scores, and dealt with simple problems in Geometry. The scores for these exercises are shown in Table VI.

TABLE VI.—SCORES SHOWING THE COMPARATIVE VARIABILITY IN GEOMETRIC ABILITIES

Exercise	Sex	Number	Mean	Standard deviation	Coefficient of variability
9	Males	1289	5.9	2.6	.44
	Females	1504	4.7	2.3	.51
11	Males	924	8.3	3.1	.37
	Females	921	7.3	3.5	.48

The conclusion Touton draws is that;

"In the complex trait called *ability to solve geometrical originals*, the median boy does excel the median girl." Our interest is in the fact that here, again, where the median is higher for the males, the coefficient of variation is shown to be higher for the girls. The

standard deviation shows the boys to be more variable in Exercise 9 and the girls to be more variable in Exercise 11.

In 1920 Whitmire[48] gave the National Intelligence tests to 1340 unselected children to study the sex differences within the group. Incidentally the differences in variability were also computed. Although the numbers for any one age are not large enough to be conclusive in a study of variability, they are much larger than many of the studies reported in the history of the discussion of this hypothesis. Since age 10 was considered in one of the former studies we shall examine the scores for this age together with those for age 11. The scores are shown in Table VII.

TABLE VII.—SCORES SHOWING THE COMPARATIVE VARIABILITY IN PERFORMANCE OF THE NATIONAL INTELLIGENCE TESTS

Age	Sex	Number	Mean	Standard deviation	Coefficient of variation
10	Males	95	148.5	52.	.35
10	Females	97	158.1	44.8	.28
11	Males	99	171.9	51.	.29
11	Females	98	190.	54.	.28

At the conclusion of this study the author states that the results show a marked tendency toward the greater variability of boys. (It should be added that the Standard Deviation was higher for the boys at most ages.) For the two ages shown above we see again the characteristic relationship. The girls show a higher mean performance at both ages. At age 10 the SD was greater for the boys, and at age 11 it was greater for the girls. The coefficient of variation is larger for the boys in both cases because of the higher mean of the girls. Variability seems to shift from age to age as well as from problem to problem when sex comparisons are made.

One other study will be given consideration because it appears to be the only one in recent years to dignify this subject with a title all by itself. Frazier[12] made a comparative study of the variability of 62,219 boys and girls 13 years of age on the basis of grade location. He concluded that the greater variability claimed for boys does not exist. The coefficient of variation for the boys was 16 and for the girls 14. Judging intelligence by grade location would scarcely justify his empirical conclusions. An educational ideal seems to have

instigated such a research rather than facts. Grade location does not correlate as well as might be desired with intelligence. From the study of "Gifted Children"[38] quoted above, it was shown that the gifted child was 48 per cent of his age above the normal in intelligence and only 14 per cent of his age above the normal in grade location. Furthermore the selection of the age of 13 in any study of variability might be unfortunate because of the effect of the adolescent advancement of females over males at this age. It is of interest to note, however that there appears to be little difference in variability in grade location.

The facts reviewed from comparative studies of males and females in the intellectual field as well as the anatomical field seem to warrant the following conclusions relative to the hypothesis of variability.

1. Where a large population of a given age has been tested the difference in variability is practically negligible.

2. Males have been found to be more variable when some characteristics or types of performance are measured and females have been found to be more variable when other characteristics or types of performance have been measured.

3. Reported differences in variability might have been due to peculiarities of formula or selection of formula for determining variability. Measures of gross variability have often shown one result while measures of relative variability have shown the reverse result.

4. Mean differences in the abilities of the sexes in various kinds of performance seem to have been the determining factor when the coefficient of variation was used as the measure of variability. Since girls scored higher than boys in most of the tests used in studies of variability, this formula anomaly has made it appear that boys were more variable.

5. Where total range is used as a measure of variability the females show a spread equal to that of males. No justification for the occasional practice of referring to the "law of greater male variability"[33] as assurance of the trustworthiness of data has been found, neither does there appear to be justification for the "restrictions of women to the mediocre grades of ability"[42] because of the greater variability of men.

Bibliography

1. Berliner, A.: Æsthetic Judgments of School Children. *Jour. of Applied Psych.* 1918, p. 229.
2. Bobertag, O.: Dritter deutscher Kongress für Jugendbildung und Jugendkunde zu Breslau 1913. *Zsch. f. angew. Psycholie*, Vol. VIII, 1913, 345.

3. Bonser: The Reasoning Ability of Children. *Amer. Jour. of Sociology*, 1910, p. 20.
4. Brown, W.: Correlation of Mental Abilities. *British Jour. of Psych.*, 1910, p. 296.
5. Castle, Cora S.: Statistical Study of Eminent Women. *Amer. Jour. of Sociology*, 1913.
6. Cattell, J. McK.: Statistical Study of Eminent Men. *Popular Science Monthly*, Vol. LXII.
7. Courtis, S. A.: "Report on the Courtis Tests in Arithmetic." N. Y. Com. on School Inquiry, Vol. I, 1911, p. 391.
8. Ellis, Havelock: "Man and Woman." Fifth Edition, p. 412 (Appendix).
9. Frailey, L. E. and C. M. Crain: Correlation of Excellence in Different School Subjects. *Jour. of Ed. Psych.*, 1914, Vol. V, p. 15.
10. Frazier, F. W.: *A Statistical Study of Sex Differences in Intelligence*, 1918. Masters Thesis, Stanford Library.
11. Fox and Thorndike: "Sex Differences in Arithmetical Ability." Columbia Contributions to Phil., Psych., and Educ., XI.
12. Frazier, G. W.: A Comparative Study of the Variability of Boys and Girls. *Jour. of App. Psych.*, Vol. III, 1919, p. 151.
13. Goddard, H. H.: Two Thousand Normal Children Measured by the Binet Measuring Scale of Intelligence. *Amer. Jour. of Psych.*, Vol. XVIII, 1911, p. 232.
14. Henman, V. A. C. and Livingston: Comparative Variability at Different Ages. *Jour. of Educ. Psych.*, Vol. XIII, 1922, p. 17.
15. Hollingworth, L. S.: The Frequency of Amentia as Related to Sex. *Medical Record*, 1913.
16. Hollingworth, L. S.: Variability as Related to Sex Differences in Achievement. *Amer. Jour. of Sociology*, Vol. XIX, p. 510.
17. Hollingworth, L. S.: Comparison of the Sexes in Mental Traits. *Psychological Bulletin*, Vol. XVI, 1918, p. 427.
18. Hollingworth, L. S. and Montague, Helen: The Comparative Variability of the Sexes at Birth. *Amer. Jour. of Psych.*, Vol. XX, p. 335.
19. Hollingworth, H. L.: "Vocational Psychology." Page 232.
20. Hollingworth, H. L.: Judgments of Persuasiveness. *Psychol. Bulletin*, 1913, Vol. XXVIII, p. 4.
21. Hardlička, A.: Anthropological Investigation of 1000 children. *Amer. Jour. of Sociology*, 1899.
22. Keyes, C. H.: Progress through the Grades of City Schools. *Amer. Jour. of Sociology*, Vol. XVII, p. 15.
23. Kuper, Gertrude: Group Differences in the Interests of Children. *Jour. of Phil., Psych., and Scientific Method.* 1912, p. 377.
24. Meyers, Garry C.: Incidental Memory. *Am. Jour. of Sociology*, 1913.
25. Munsterberg, Hugo: "General Psychology." 1914.
26. Mulhall, E. F.: Tests of the Memories of School Children. *Jour. of Educ. Psy.*, Vol. VIII, 1917, p. 295.
27. Patrick, G. T.: In Quest of the Alcohol Motive. *Popular Science Monthly*, September, 1913.
28. Pearson, Karl: Chances of Death. *Amer. Jour. of Sociology*, 1897.

29. Pfarmkuch, W.: Ueber die Köyurform der Neugeborenen. *Arch. f. Gynak.* Bd. IV.
30. Pressey, Luella W.: Sex Differences Shown by 2544 School Children. *Jour. of Applied Psych.*, Vol. II, 1918, p. 323.
31. Pyle, W. H.: Sex Differences and Variability. *School and Society*, 1918, Vol. VIII.
32. Race, H. V.: A Study of a Class of Children of Superior Intelligence. *Jour. of Educ. Psych.*, 1918, p. 91.
33. Rogers and McIntire: Intelligence of Children. *Brit. Jour. of Psych.*, Vol. VII, 1914, p. 265.
34. Stone: Arithmetical Abilities. *Amer. Jour. of Sociology*, 1908, p. 36.
35. Strong, E. K.: Relative Merits of Advertisements. *Amer. Jour. of Sociology*, 1911, p. 78.
36. Terman, L. M.: "The Measurement of Intelligence." Page 70.
37. Terman, L. M. and others: The Stanford Revision and Extension of the Binet-Simon Tests for Measuring Intelligence, 1917, Chap. IV.
38. Terman, L. M. et al.: "Genetic Studies of Genius." Vol. I, 1925, Chap. III, p. 39.
39. Thompson, H. B.: The Mental Traits of Sex. *Amer. Jour. of Sociology*, 1906.
40. Thorndike, E. L.: "Principles of Teaching." Page 96.
41. Thorndike, E. L.: "Educational Psychology." 1910, p. 35, 42.
42. Thorndike, E. L.: "Sex in Education." The Bookman, Vol. XXIII, p. 213.
43. Thorndike, E. L.: "Educational Psychology." Vol. III, 1914.
44. Thurstone, L. L.: Variability in Learning. *Psych. Bulletin*, Vol. XV, 1918, p. 210.
45. Touton, Frank C.: Sex Differences in Geometric Abilities. *Jour. of Educ. Psych.*, Vol. XV, 1924.
46. Wells, F. L.: Sex Differences in the Tapping Tests. *Amer. Jour. of Psych.*, 1909.
47. Wells, F. L.: Relation of Practice to Individual Differences. *Amer. Jour. of Psych.*, 1912.
48. Whitmire, Ethel Dorothy: "A Study of Sex Differences in 1340 Unselected Children." Masters Thesis, Stanford Library, 1920.
49. Wooley, H. T.: The Psychology of Sex. *Psych. Bulletin*, Vol. XIX, 1914, p. 351.

AN EXPERIMENT IN MOTIVATION

CLAY CAMPBELL ROSS

University of Kentucky

It has been about half a century now since psychology broke away from its family roof, philosophy, and definitely set up in business for itself as an experimental science. During this period the field of inquiry has been extended from a mere analysis of consciousness to the study of adaptive behavior in general, and now includes the physical reactions as well as the conscious. Psychologists have explored and charted the acquisition of skills in many directions. We have learning curves galore, both for animal and for human subjects.

Not satisfied with this achievement psychologists have turned their attention to the nature of the capacity which makes learning possible. And so we have the many attempts to measure innate intelligence, or whatever it is that lies back of learning and sets the limits for its development.

But a problem that lies back of even this has received but scant attention. Behind the *act* of learning is the *capacity* to learn, and back of the *capacity* is the *motive* to learn—the desire, urge, impulse, drive, or something, that makes the creature *want* to learn, that pushes him out to meet his environment. One of the reasons why most of the correlations of mental capacity with actual achievement, in school and out, have been disappointingly low, has been that students of real ability have not felt a proper urge to work, while those of mediocre talent have frequently possessed the urge to achieve.

If we are ever to be successful in our efforts to predict achievement, therefore, we must not be content with merely analyzing the learning process, understanding the mechanism of learning, its structure and laws of operation; nor with merely exploring the height and range of human possibilities; but we must also find out about the dynamic aspects of human nature. We must discover not only *how* the mind works, but *why* it works when it does and the way it does. In fact, just now the "Why" is more important than the "How."

And yet it is just at this point that psychology has made the least convincing experimental approach and still relies most heavily upon philosophy. Nor is the reason far to seek. It is inherently a most difficult problem. Behavior, both of the conscious and the physical types, can be observed directly, but how can we look behind our behavior to see what makes us behave the way we do? Obviously,

the motive can only be inferred rather than observed directly. But this is also true of intelligence. It, too, can only be inferred from the behavior observed. And so the problem of motivation, while difficult, seems not an impossible one. But it must be admitted, I think, that our knowledge of motivation, despite its fundamental importance, is at present most inadequate and unsatisfactory.

The present experiment attempts to inquire briefly into one aspect of this problem, namely, the relation between knowledge of progress and achievement, in one form of motor learning. In other words, to what extent does knowledge influence motivation or, in terms of the older psychology, to what extent does cognition influence conation, insofar as this is reflected in the results accomplished?

The process involved was a simple act of muscular skill, namely, making groups of four vertical lines and crossing them with a fifth, after the order of the device used in tabulating, sometimes called "tallying." The subjects were 59 college students in psychology, who recited in four small groups, of whom 28 were freshmen and 31 were upper-classmen.[1] All of these students were more or less familiar with the general theory of motivation and the influence of various factors upon the learning curve, and presumably had an interest in the problem.

The experiment was conducted in the following manner: An initial practice period of one minute was given each of the four groups. The practice had been preceded by uniform instructions as to what they were to do, and they were told to do their best, but to avoid keeping track of the progress made. The papers were collected immediately following the practice period. Later they were scored by the experimenter and each group was divided into three sections of equal size and of substantially equal achievement on the first practice period, care being taken not only to keep the averages equal but to balance off an especially good student in one group with one equally good in the other two groups, and so on, insofar as this was possible.

At the beginning of each of the succeeding practice periods the following procedure was observed: One of the sections in each of the four groups was given full information as to its progress from day to day, each student being shown his paper for the preceding practice period with corrections and score indicated thereon, and a frequency table for that section was put on the board. Each student in this

[1] Miss Nira Klise, Instructor in Psychology, Iowa State College, assisted in collecting, tabulating, and analyzing the data.

section was urged to watch his relative and absolute progress from day to day and make a learning curve.

A second section in each group had partial knowledge of progress. That is, they were told who were above average and who were below average in that particular section, although they did not know how much above or below average they were or indeed, whether it was the good, poor, or average section to which they belonged. However, they did know a little as to their relative achievement in their own section, and that was all. Obviously the status of their knowledge was quite vague and unsatisfactory.

The third section in each group had no information whatsoever as to their own progress, although they heard each time what was said to the second section, and everybody knew what the first section was doing. Along with the rest, they were urged each time to do their best and to refrain from keeping track of their progress. To make it easier to remain in ignorance as to progress being made, paper of a different size and shape was provided for each practice.

The purpose of this experiment, therefore, was to compare the achievement of the three sections in each of the four groups, one section working with full knowledge of progress, a second section with partial knowledge of progress, and the third with no knowledge of their own progress at all. The object of this comparison was to discover to what extent a knowledge of results is a factor in motivation.

The groups met every other day, and no practice was allowed in between. At the beginning of each recitation period the first section looked over their papers and noted progress on the last practice period, the second section was told who were above and who were below average, and the third section was told nothing at all. Then all were given one minute's practice, and all were told to do their best. This procedure was followed for nine successive practice periods following the initial period.

On the eleventh day the process was reversed. That is, the section in each group which had not known their progress, and the one which had had only partial knowledge of progress, were shown their papers for the tenth practice period, and each pupil in these sections was given his individual record for the ten periods, together with the averages for all the sections. On the other hand the section in each group which had kept track of their progress from day to day was not shown their achievement for the tenth practice period, although the members of it did learn for the first time how they had been standing

with reference to the other sections. After the first two sections had had time to study the record of their progress so far, two additional practice periods of one minute were held. After the first of these periods the sections which had been working hitherto with no knowledge of results and with partial knowledge, were given time to determine their scores and a frequency table showing the results for each section was put on the board, while the section which had been working with full knowledge of results during the first nine periods was not permitted to look at their papers at all. Then a final practice period was given and all papers were scored.

TABLE I.—AVERAGE SCORES FOR THE THREE SECTIONS. FOUR GROUPS COMBINED

Trial	Corrected scores			Uncorrected scores		
	Amount of information			Amount of information		
	Full	Partial	None	Full	Partial	None
1	35.6	35.7	35.0	38.5	37.3	37.7
2	44.8	45.2	44.7	45.6	45.5	45.8
3	50.5	47.8	46.6	50.4	48.1	48.8
4	51.4	50.3	49.0	51.9	51.0	50.8
5	53.4	50.1	51.1	53.8	51.4	52.9
6	56.2	52.0	52.0	56.9	54.2	54.1
7	54.9	51.6	51.7	55.4	53.0	53.4
8	57.2	52.6	50.8	57.3	54.1	52.4
9	57.7	53.8	51.8	58.7	54.7	53.9
10	57.4	54.3	53.2	57.7	55.2	53.9
	None	Full	Full	None	Full	Full
11	58.7	56.0	54.7	59.1	56.5	55.3
12	61.6	58.1	56.5	61.7	58.7	56.8

The results for the four groups combined are given in Table I and in Fig. 1. It will be noted in Fig. 1 that all sections start off equal on the first practice period, and that the progress was practically identical for the three sections on the second practice period, but that after that the effect of novelty wore off, the section knowing its progress forging steadily ahead, while there was but little difference between the section with partial knowledge and the section with no knowledge, although in general the one with no knowledge at all was a little behind the other. The superiority of the section working with full information as to progress may be summarized as follows: From the first

practice period to the tenth inclusive, the section with full information gained from 2.2 per cent to 8.5 per cent more than the section with only partial information, the average advantage being 6.2 per cent, while the section with full information gained from 4.5 per cent to 12.6 per cent more than the section with no information, the average advantage being 8 per cent. These same characteristics are likewise apparent in the curves for the individual groups, their form being essentially the same in every case.

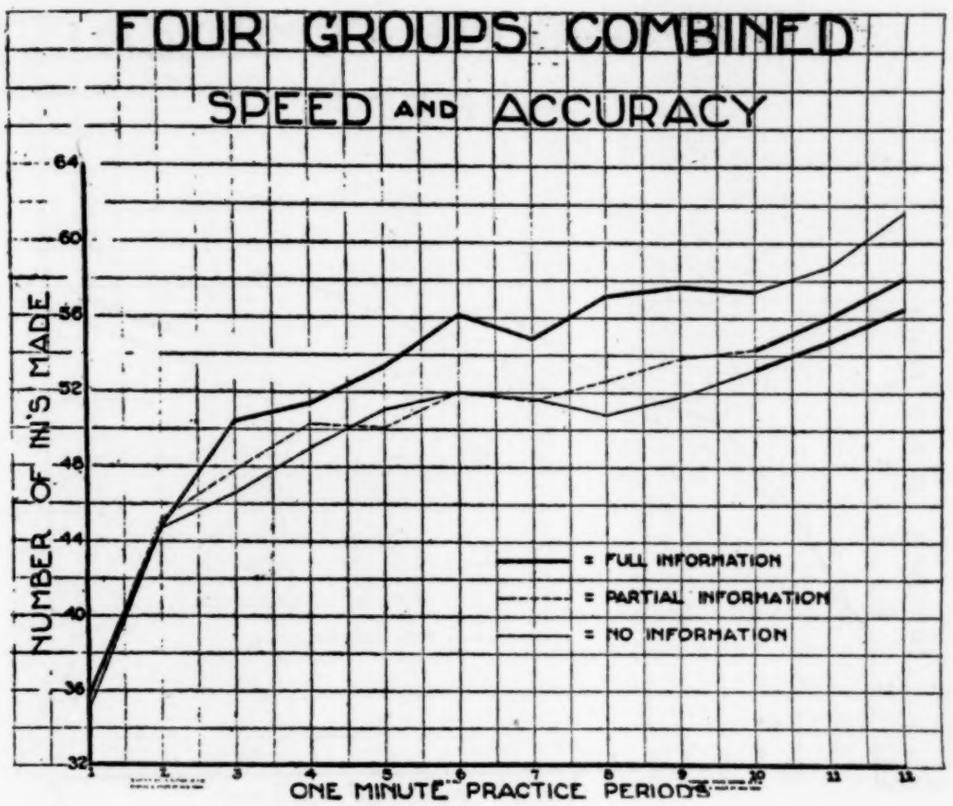


FIG. 1.

The results so far considered are in terms of corrected scores, certain directions for scoring being observed. For example, no credit was allowed for a unit of work if the line crossing the four vertical lines failed to touch all of the four, or where there were fewer than four lines or more than four lines in the group crossed. These scoring directions were explained at the beginning of the experiment but were not referred to again, however, the group working with full information could see the ones checked off as errors each time, and so had the advantage of

this knowledge over the other two sections, in addition to the knowledge of the total scores from day to day. The results so far considered, therefore, are in terms both of speed and accuracy.

Now let us turn to a consideration of speed only. In calculating this scoring directions were disregarded, any five lines being considered a unit, whether properly crossed or not. Figure 2 shows the results for the four groups combined. When scored on this basis the section with full information continued to have the advantage, but the margin

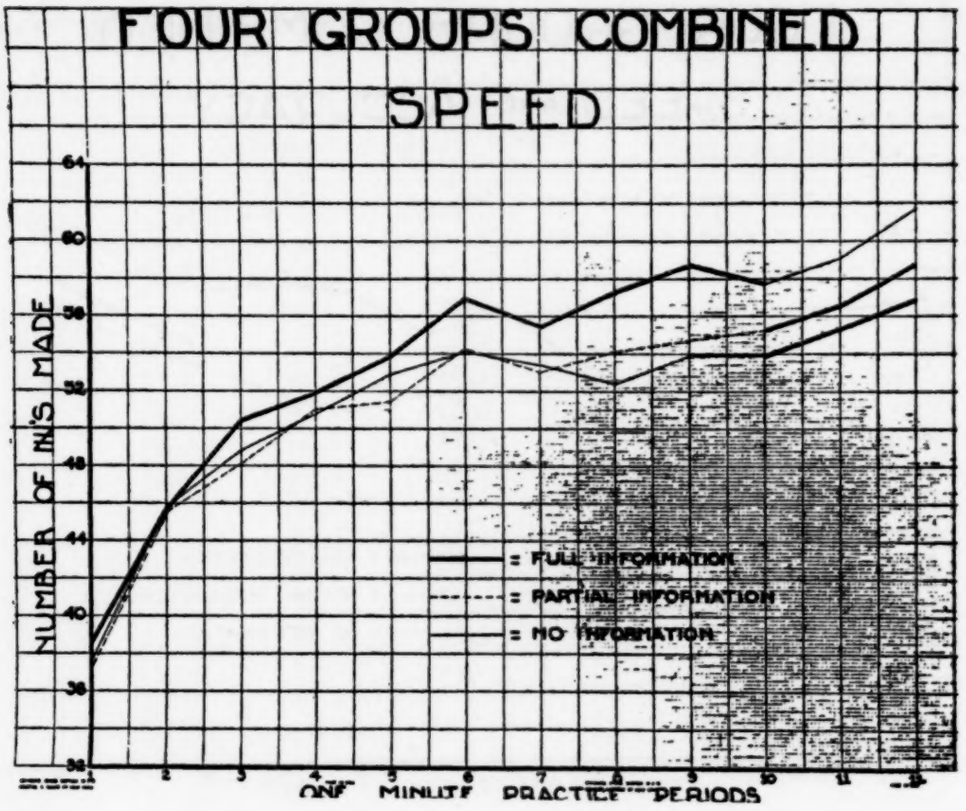


FIG. 2.

was somewhat reduced, and it is clear that a part of the advantage shown in Fig. 1 was due to a knowledge of the nature and extent of the errors made, as well as to a knowledge of the total scores on each trial. The pedagogical implications of this, of course, are perfectly obvious and quite important.

The results on the last two practice periods, when conditions were reversed, deserve special attention. An interesting comparison is possible with Book's experiment in making legible "a's," reported in

the *"Pedagogical Seminary"* for December, 1922. The practice periods were a half-minute in length for ten trials, and then the conditions were reversed, much as in the present experiment. Book's results are given in Fig. 3. The curves for the first ten trials are quite similar to those found in Fig. 1 above, but after the conditions were reversed, the curves are quite different in the two experiments. Book found that the group working with full information lost their advantage as soon as the information was withheld. That is, the impetus

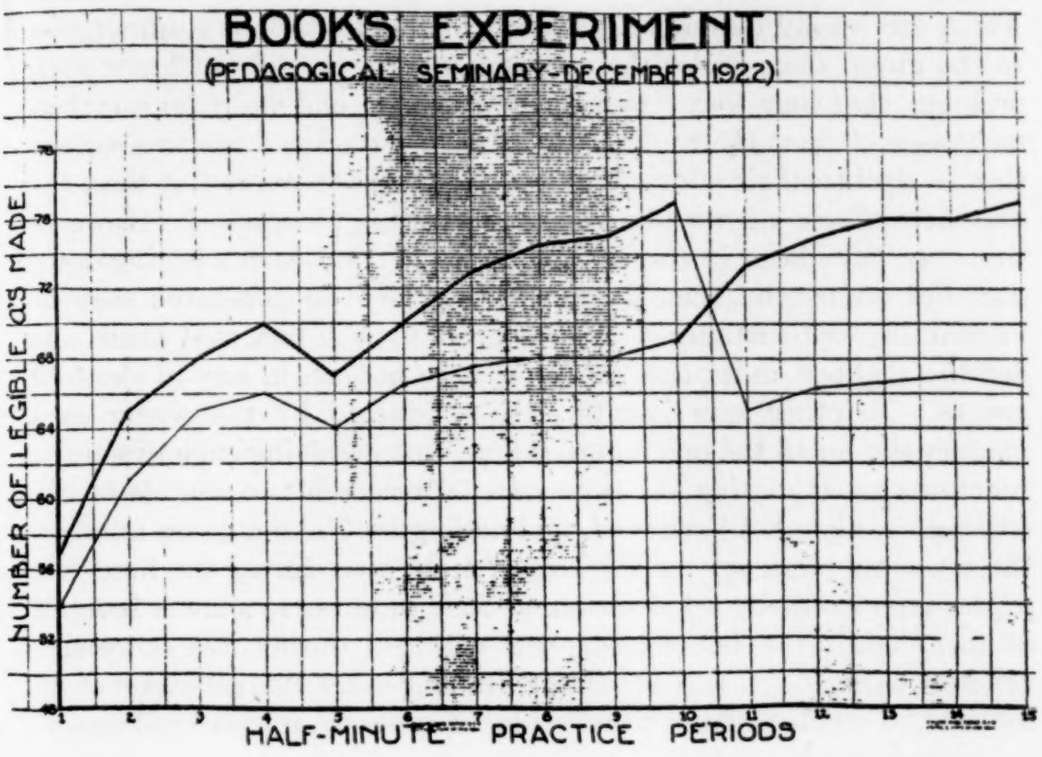


FIG. 3.

of the drive did not persist beyond the immediate practice period where the information was given. In fact, the drop in the curve is so sudden as to suggest that found in the after-effects of powerful artificial stimulation such as is afforded by drugs. But in the present experiment the influence of the drive persists undiminished through the last two periods. It appears that the momentum gained during the first ten practice periods is sufficient to carry the students over the two remaining periods, at least. Clearly this is as we would like to have it. In fact, one may well question the virtue of a motive which is so temporary

in character as Book found, and especially one that is followed by such immediate and unfavorable after-effects. On the other hand, it is very much worth while to find a motive whose effects are both immediate and lasting, and it is most desirable of all to discover a motive that, in addition to these two characteristics, is self-sustaining when once aroused, no matter what its origin. Nor is such a form of motivation an unreasonable expectation. It is notoriously true of anger, for example, that when once aroused it is likely to persist for some time and greatly influence our subsequent behavior, even in situations which are wholly dissimilar. May it not be that Book's admonition to the group that, under the reversed conditions, did not know their progress, that they were "to banish all thought and desire for improvement as such from their minds," may have served as a positive suggestion to slacken their efforts that was by far more important than the negative effects of withdrawing information of results? However this may have been in the Book experiment, that such a letting down *need* not occur, and that the motivating force so generated may be measurably self-sustaining, is evidenced by the fact that there was not the slightest indication at all for it to happen in any of the four groups. In other experiments since conducted by the writer in a manner similar to the one reported here, but involving such dissimilar functions as estimating the time interval between two sounds heard, writing the alphabet backward, and adding by 7's to a given number, there was no tendency for the groups motivated during the first part of the experiment by a knowledge of their progress to show a reversal of form when the information was withheld during five successive practice periods. If this is found to be a general characteristic of all motivation resulting from a knowledge of progress, its practical importance is very great indeed.

Which pupils make the most rapid progress, those with high initial scores, or those with low initial scores? And with which type of pupil is the effect of motivation greatest? To answer these questions the pupil in each section of the four groups who made the best score on the first two trials was selected for special study; and likewise, the one from each section who made the lowest score on the first two trials. The average score of the four "best pupils," so selected, and of the four "poorest pupils," so selected, are given in Figure 4. The results may be summarized as follows: The rate of improvement for all sections is greater for the "poorest group" than for the corresponding sections in the "best group," but the rate of superiority for the full

information section over the other two sections is about the same for the two groups.

In a similar manner a "brightest group" and a "dullest group" were formed by selecting the brightest and dullest pupil in each section, as measured by a group intelligence test. The influence of motivation was greater for the bright pupils than for the dull. In fact, while the very best record of all was made by the bright pupils working with full information, the poorest record of all was made by

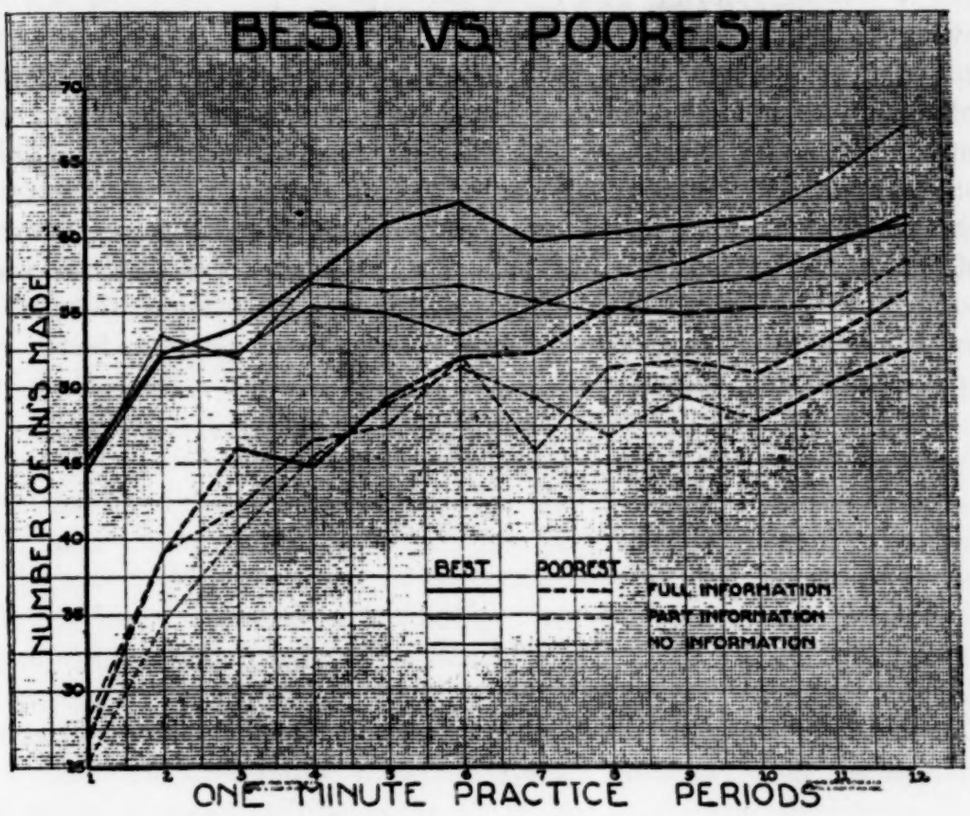


FIG. 4.

the bright pupils working with no information as to progress. When all pupils were considered, however, there was found to be but slight correlation between general intelligence and motor ability of this particular type at any stage of the practice.

Just a word in conclusion. In interpreting these results it should be kept in mind that all these pupils were students of motivation and presumably were interested in the problem, that they knew something of the general effects of a knowledge of results upon achieve-

ment and all knew that one group was making improvement during the progress of this particular experiment, and that there is no reason to believe that any of these factors operated more decisively in one section than in another. Moreover, all were equally subjected to the facilitating effects of hearing each other's pencils scrape upon the paper during practice, and the like. Insofar as could be determined the one factor that was variable was the knowledge of one's individual progress, both relative and absolute, and that of the section of which was one a member, during this particular experiment. In other words, even in highly motivated groups, as undoubtedly these were, the addition of a single other motivating factor, namely, knowledge or results, is sufficient to give the pupils with such knowledge a distinct superiority over the others, and the degree of superiority is roughly proportional to the amount of information possessed. To what extent this result is to be explained by the hypothesis that additional sources of energy were released in the subjects who knew their progress, due to the mild emotional effects accompanying rivalry, either social rivalry or self-rivalry, is not known, but this was probably a factor of considerable importance.

NOTES AND COMMUNICATIONS

A Further Note on the Use of the Spearman Prophecy Formula: A Correction

In the February[1] issue of this Journal appeared a note on the use of the Spearman Prophecy Formula, pointing out precautions which should be taken in using it. Published data by Ruch, Ackerson, and Jackson were used in illustration.[2]

Unfortunately, in the computational work on their paper an error was made which increased the difference found between their average correlation at the longer interval, and that at the shorter interval, (*loc. cit.*, p. 310) so that it appeared to be over 2 PE. Actually, if correctly computed, the difference almost disappears. It was, in view of its smallness used merely to indicate the "possible presence" of a regular variation. It may be said, however, that while theoretically such a difference might be expected, it is not present in the published data of Ruch, Ackerson, and Jackson.

Correspondence has passed between Professor Ruch and myself on the subject. Professor Ruch says: "I believe that you have called attention to a point which we overlooked in choosing our 20 correlations. I am willing to grant the possibility you have raised about a systematic decrease in the correlations with time. [But] none of the possible ways of examining our published data shows any real evidence that it is subject to such effects (although, as stated above [I] hold them to be the usual finding). It may be that we were merely lucky in avoiding systematic effects, [but] spelling is, as is well known, notoriously refractory to practice in the sense we commonly observe practice in duplicate forms of educational and mental tests."

The over-prediction found by Ruch, Ackerson, and Jackson is, however, "statistically significant" in the case of the early tests, and though the data published in their article do not reveal the cause, there must be one. It would be interesting to know whether even in spelling slow, systematic changes do not take place, which have sufficient total effect to produce the overprediction found.

It is regretted that an arithmetic error should have led to an implied criticism of the work of Ruch, Ackerson, and Jackson, par-

[1] C. S. Slocombe: The Spearman Prophecy Formula. *Journal Educational Psychology*, 1927, p. 125.

[2] Ruch, Ackerson, and Jackson: An Empirical Study of the Spearman-Brown Formula. *Journal Educational Psychology*, 1926, p. 309.

ticularly in view of the fact that their published data indicated that much care was taken to show that all variables had been eliminated.

C. S. SCOLOMBE.

Lincoln School of Teachers College.

COMMUNICATIONS

THE INTERNATIONAL CONGRESS OF PSYCHOLOGY

The Eighth International Congress of Psychology was held at Groningen, Holland, during the week of September 6 to 11. The meetings were held in the University of Groningen. About 300 members were registered, including psychologists from the principal countries of Europe and America. Professor Heymans of Groningen acted as chairman of the Congress. All arrangements for the entertainment of the visiting psychologists were excellent and the hospitality of the people of Groningen was bounteous.

The papers and discussions in general impressed the American psychologist as being very theoretical, many, indeed, very philosophical. There was a marked absence of reports on experimental work recently done, as contrasted with a tendency to set up or defend a theory or point-of-view.

Educational psychology, as we know it in America, was very sparsely represented. Only one paper dealt with the psychology of learning. This was a paper by Decroly (Brussels) describing the advantages of teaching reading and writing by the "global" method, by starting with whole phrases instead of with smaller units. With this method he has been able to teach deaf children of five and six to read and write simple phrases after a few months of instruction.

In the field of tests there were five papers. Buyse (Tournai) reported on the results of Myers Mental Measure and Ballard's Tests with several thousand children. He showed differences between classes, schools and neighborhoods and argued for better classification. Buyse also reported on the use of tests for the selection of gifted children for higher education in one of the districts of Brussels. He described the different tests used in 1923, 1924 and 1925. The last test was a combination of concrete and abstract items, working time three hours. It proved satisfactory for selecting the gifted. Decroly (Brussels) gave a brief report of an extensive intelligence survey in Belgium covering 7201 pupils with a translation of the

Ballard Test. The results showed interesting differences between the two language groups in Belgium, as well as differences between urban and rural, industrial and non-industrial districts. Very slight sex differences were apparent. The Belgian age standards fall below Ballard's after age 10. From age 11 to 14 where we find the largest numbers of Belgian children, the norms are lower than Ballard's. The means for the grades are almost identical for the English and Belgian children. Foucault (Montpellier), reported on the measurement of rapidity of movement using six individual tests with a small number of children at each age from three up to the adult level. Jones (Leeds) reported some results with the Seashore Musical Ability Tests. He found that the Consonance Test gave widely different results for the same individual when tested at different times. He also calculated what percentage of votes agreed with the Seashore Key and found several discrepancies, due to the difficulty of knowing precisely what to judge. In the realm of child psychology Charlotte Buhler (Vienna) described her observations of the social reactions of young children with reference to other children.

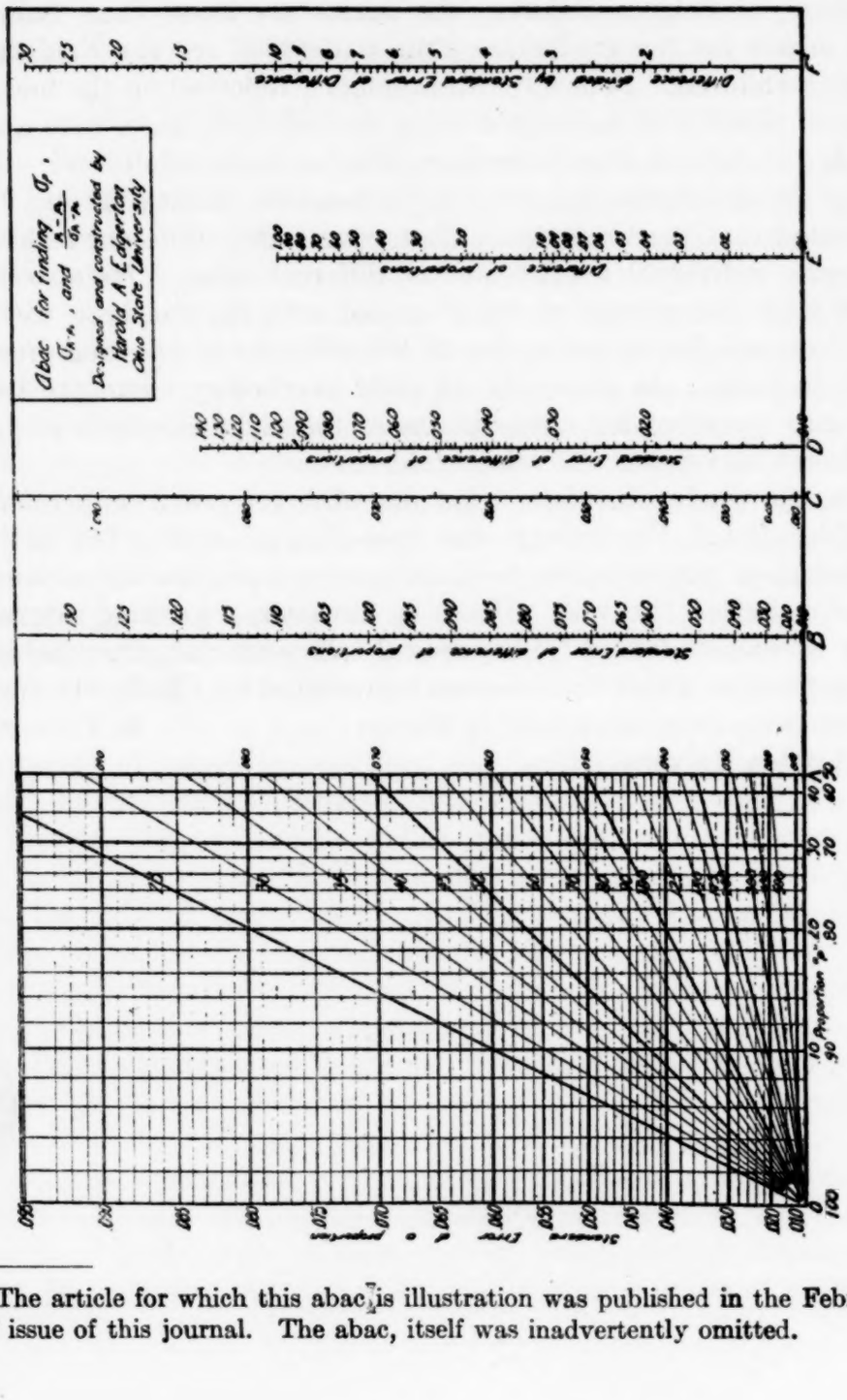
On the whole, therefore, what one observer noted with reference to Educational Psychology was exceedingly small. The attitude of European psychologists towards tests is much like the attitude in America before the war. There is, however, a growing interest in them as was evidenced from personal talk with many psychologists. My impression is that the Congress represented very badly the amount of work with tests being done in Europe.

R. PINTNER.
Columbia University.

AN ABAC FOR FINDING THE STANDARD ERROR OF A PROPORTION AND THE STANDARD ERROR OF THE DIFFERENCE OF PROPORTIONS[1]

HAROLD A. EDGERTON

Ohio State University



[1] The article for which this abac is illustration was published in the February, 1927 issue of this journal. The abac, itself was inadvertently omitted.

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES

REPORTED BY C. O. MATHEWS
Research Associate, Institute of Educational Research, Teachers College
Columbia University

Psychology of Learning

The Transfer of "Translation Thinking." D. R. Smith, A. R. Mead and C. C. Peters. School and Society, Mar. 26, 1927, 380–384. The results with groups of pupils carefully paired show that the transfer from ability to translate Latin to ability to reason is very slight.

Consciousness of Method as a Means of Transfer of Training. G. P. Meredith. The Forum of Education, Feb., 1927, 27–45. The study was planned to test which of two different methods of teaching science was the more effective in securing transfer from training in defining scientific terms to defining ordinary terms in daily use.

An Experiment in Kinesthetic Learning. Charles L. Bowdlear. American Physical Education Review, Feb., 1927, 100–105. A report is given of the progress of learning a certain method of mounting parallel bars from the records of 37 high school pupils.

Early Emotions and Early Reactions as Related to Mature Character. Charles H. Judd. School and Society, Mar. 26, 1927, 355–360. The effect of training during early childhood upon the development of desirable attitudes is pointed out.

Psychology of School Subjects

The Superiority of Distributed Practice in Drill in Arithmetic. F. B. Knight. Journal of Educational Research, Mar., 1927, 157–165. In a carefully controlled experiment materials providing for distributed practice showed significantly better results than those obtained by the use of materials which ignored careful distribution of combinations.

A Report of a Controlled Study of the Learning of Multiplication by Third Grade Children. John Guy Fowlkes. Journal of Educational Research, Mar., 1927, 181–189. One hundred combinations were taught by printed directions and their relative difficulties determined when the frequency of drill was kept constant.

Summary of Reading Investigation (July 1, 1925 to June 30, 1926) II. William S. Gray. The Elementary School Journal, Mar., 1927, 495–510. This is a continuation of previous summaries and numerous recent studies are briefly reviewed under 17 headings.

Reading Distance and Some of Its Implications. E. B. Notson. School and Society, Mar. 19, 1927, 353–354. The distance at which pupils hold books from their eyes in reading silently gradually increases from 8.2 inches in the first grade to 13.8 inches in the university.

The Effect of Summer Vacation on Silent Reading Ability in Intermediate Grades. Harry E. Elder. The Elementary School Journal, Mar., 1927, 451–546. When records of 203 third, fourth, fifth and sixth-grade pupils were combined it is shown that 59 per cent improved in silent reading during the summer, 27 per cent were less efficient and 15 per cent scored the same.

Selected Bibliography on Children's Vocabularies. Madeline Darrough Horn. Childhood Education, Mar., 1927, 316–319. The references are classified under the following heads: Age groups; school groups; vocabulary tests; relation to intelligence; distribution of parts of speech; sentence structure; theoretical discussion.

A Comparative Study of Different Methods Used in Teaching Beginners to Write. Oscar Edward Hertzberg. The Teachers Journal and Abstract, Mar., 1927, 215–219. The results of a comparative study of four different methods of teaching kindergarten children to learn to write are summarized.

Exceptional Children

Portrait Studies of Some Exceptional Children. Eve Macauley. The Forum of Education, Feb., 1927, 27–36. This is a report of case studies of four pairs of children, one pair under each of the following headings: "Imaginary Companions," "Troublesome Children," "Backward Boys" and "Only Children."

The Gifted Child. Dortha Williams Jensen. Journal of Educational Research, Mar., 1927, 198–206. Questionnaire reports of present practices in special classes for the gifted are summarized.

School Marks

"A Committee Report, College Marking Systems." Shelton Phelps. The High School Quarterly, Jan., 1927, 80–85. This report of marks given by members of the Southern Association of Colleges and Secondary Schools shows that colleges are grading too high and suggests remedial measures.

An Index of Consistency for College Students. E. B. Mersereau. School and Society, Mar. 19, 1927, 352–353. An index of consistency was obtained by correlating the highest and lowest marks for each college class. The data show that these groups became less consistent as they became more mature.

The Effect of Handwriting upon Grading. H. W. James. The English Journal, Mar., 1927, 180–185. Good handwriting is shown to influence grade to the extent of 8.63 points on a scale of 100 points in favor of good writing.

Miscellaneous

The Technique of Teaching Educational Measurements in Normal Schools. Helen Gay Pratt. Educational Administration and Supervision, Mar., 1927, 200–204. This is a discussion of the principles upon which the successful teaching of measurements depends.

The Overlapping of Subject Matter in Courses in Education. Richard E. Hyde. The Journal of Educational Method, Mar., 1927, 306–309. "A detailed study of ten texts in education, used in different courses, shows an overlapping of subject matter among them amounting to a total of 51.6 per cent of the contents."

The Play Interests of Children of Ages Five to Eight. Harvey C. Lehman and Paul A. Witty. Childhood Education, Mar., 1927, 299–303. By the use of a play quiz the authors determined the most popular play activities of children of each age.

The Method of Paired Comparisons for Social Values. L. L. Thurston. The Journal of Abnormal and Social Psychology, Jan.-Mar., 1927, 384–400. A method of constructing a scale of values based upon judgments is explained in detail.

The Detection of Cheating in Objective Examinations. Charles Bird. School and Society, Feb. 26, 1927, 261–262. A comparison of identical errors made by members of a class furnish data for the detection of cheating.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

CONDUCTED BY JOHN HOCKETT
The Lincoln School of Teachers College

A General Survey of Mental Development

Mental Growth and Decline, by H. L. Hollingworth. New York: D. Appleton & Co., 1927. Pp. VIII + 396. $3.00.

This "volume originated as a series of college lectures, which attempted to survey our knowledge of individual mental development." It remains such a survey, but does not limit itself to mental development only. It discusses at some length other factors, which perhaps influence mental growth, and also those products of human endeavor which constitute so-called social heredity. The consideration is of the development of individuals, commencing a few generations before they are born, and concluding a few decades after they have died. Only 100 pages are devoted to stages included within the school age, though these are fairly well covered. (Is this an indication of the paucity of knowledge regarding children really possessed by educators?)

The treatment is general, elementary and all-inclusive, as befits a book apparently intended for college freshmen. It touches only lightly on controversial matters, and does not expose fallacies in the logic of many theories; but it gives a very fair preliminary survey of the field, which might well become the background of more specialized study.

In one respect only does the term "preliminary survey" not properly describe the book; for such a survey usually consists in a general framework with a few details but roughly drawn in. Hollingworth has been overgenerous in the matter of details. The framework is thus in places somewhat obscured, by the mention of what are really matters of common knowledge. Their omission would have assisted in the prevention of ennui, which the reading of the book produces.

His style of writing is a little too wordy and monotonous—perhaps a minor consideration, if the book is to be used by college students—though this is offset somewhat by the quiet humor interspersed throughout. Hollingworth's suggestions for the utilization of grand-

parents are as original as is his placement of Shakespeare with Kilpatrick and Thorndike as authorities on the stages of development.

C. S. Slocombe.

Lincoln School of Teachers College.

Real Help for the College Student

Methods of Study, by Claude C. Crawford. Published by the author, University of Southern California, 3551 University Avenue, Los Angeles, California, 1926. Pp. 163.

I recently had the privilege of reviewing three books on the problem of study. That by C. C. Crawford is superior in many respects to any other book on the subject that I have seen. It is written for the college student but does not follow traditional lines. Crawford used a systematic method of assembling the material for this work. "The steps in the process of assembling the material were as follows: *First*, a complete difficulty analysis was made. This was a list of all difficulties that could be discovered that confronted college students in their study. *Second*, all available methods were collected for solving the respective difficulties. These were collected by actually interviewing a large number of college professors and other outstanding scholars and by getting written statements from several hundred college students as to their methods of solving the difficulties. *Third*, the methods were classified, arranged in outline form, interpreted in terms of psychology as far as possible, subjected to experimental testing in some cases, and arranged in manuscript form. This material was used in lecture form and later revised for the present publication." The result is a manual packed with detailed and practical suggestions as to methods of work. The objective, impersonal nature of the book is especially commendable. There is a minimum of pleading or exhortation to take interest in one's studies, to concentrate, to banish disturbing thoughts, to follow a daily schedule. The paragraph headings are participial rather than hortatory. Practically no space is given to developing ideals, or will power, or interest in one's work. Much space with a wealth of detailed suggestions is given to working conditions, use of text books, note taking, use of the library, writing a paper, outlining, and preparing for an examination. I have seen no better book for use in the college course on "How to Study."

As an example of the originality of the contribution of the book, take the quotations from eight college professors on the methods they employ in organizing notes for writing. Or take the following quotations in the chapter on Thinking. "Writing is a very effective ally and servant in original thinking. The record of your previous thoughts serves to prime the pump and start the flow of new ones. Writing has been compared with nails in a building to hold things together while you work in another part of the structure," or "Time permits new associations to form and new ideas to arise which cannot be forced at one sitting. When you are off duty and when you are thinking of other things this problem acts as a magnet and draws to it anything which is related to it which goes through consciousness."

Of course it must be admitted that for high school students, even the detail of this book is still too general High school students as a rule need more than being told how to study. They need to be provided with the actual study materials, to be given specific directions as to what to do in concrete situations and opportunity must be provided to develop certain habits. There is a distinction between knowing what good study is and knowing how good study is formed.

PERCIVAL M. SYMONDS.

Teachers College.

QUOTATIONS FROM THE MASTERS OF PSYCHOLOGY

Readings in Educational Psychology, Edited by Skinner, Charles E.; Gast, Ira. M. and Skinner, Harley C. New York: D. Appleton and Company, 1926. Pp. XXVII + 833.

The vast amount of available material in educational psychology as a result of the increasing number of contributions in recent years is given as the justification for this volume of readings. An attempt has been made to select references which represent diverse points of view. The extent to which this goal has been achieved may be judged partly from the following tabulation. It shows the number of selections written by various authors which are included in the book. All authors from whose publications more than two readings were selected are included. A few of the selections are by joint authors such as Turner and Betts and Norsworthy and Whitley so there is slight duplication as to the number of selections listed in the table.

Number of selections	Authors
50	Thorndike
31	Dewey
22	James
19	Colvin, Hall
18	Betts
17	Judd
14	Bagley
12	McDougall
11	Watson, Woodworth
10	Turner
9	Freeman, Keith, Pillsbury
8	Angell, Seashore
7	Book, Horne, Kilpatrick, Parker, Terman
6	Bolton, Burnham, Davenport, Dearborn
5	Guthrie, Pintner, Smith, S.
4	Dunlap, Ellwood, Guyer, Koffka, Munsterburg, Norsworthy, Redfield, Starch, Voelker, Whitley
3	Baldwin, Charters, Conklin, Whipple

Chapter I presents the Problems and Scope of Educational Psychology with 43 excerpts. It forms a good introduction to the special divisions within the field. This is followed by 23 chapters treating such topics as Components and Physical Basis of Behavior, Heredity and Environment, Intelligence, Individual Differences, Instincts, Emotions, Attitudes, Learning, Study Habits, Perception, Association, Imagination, Thinking, Attention and Interest, Play, Childhood and Adolescence, Transfer of Training, Mental Fatigue, Mental Hygiene, Interpretations in Psychology and Statistical Methods.

The longest chapter is the one on "Intelligence" covering 112 pages or nearly one-seventh of the book. The next in length is entitled "The Learning Process" covering 60 pages. However, since several other chapters also treat phases of learning more space is devoted to this subject than to any other. In the opinion of the reviewer it is scarcely justifiable to include a summary of statistical methods in such a book of readings. There are now available a number of simplified explanations of these methods.

The book is well organized. The order of topics somewhat follows that of several widely used texts. The table of contents lists each excerpt by title and author in the order printed. At the end of

each chapter one or two pages are devoted to questions and additional references. There is a nine page glossary of psychological terms and a nineteen page index at the close of the book.

The book is intended for undergraduates, and many of the selections have been adapted so as to avoid the use of highly technical terms. A number of the more simple tables and diagrams have been included in quotations. A wide use is predicted for this volume. It will make available to many beginning classes diverse points of view and will serve as a means of rapid review and survey for more advance students.

C. O. MATHEWS.

Institute of Educational Research,
The Lincoln School of Teachers College.

EDUCATIONAL AND VOCATIONAL AID FOR WOMEN STUDENTS

Guidance for College Women, by Mabelle Babcock Blake. New York: D. Appleton and Company, 1926. With an introduction by William A. Nielson. Pp. XI + 286. $2.50.

Guidance is defined by Dr. Blake as the process of assisting students to make progress in their educational and vocational careers. The need for such guidance is made apparent through the citation of a score or more of cases under headings such as, "The student who does not know how to study," "Afraid to recite," "Thought instructors unjust," "Wants to be an interior decorator."

In order to ascertain the extent to which colleges and universities are attempting to help such cases the author sent a questionnaire to 97 colleges for women and to 64 co-educational institutions. Replies were received from 51 per cent of the former and 70 per cent of the latter. They showed that about 50 per cent of these institutions hold group conferences about vocations; 71 per cent maintain placement bureaus; 60 per cent have faculty advisers; about one-third have some form of occupational counseling; more than two-thirds give a psychological test though it is doubtful that the results are made practical use of in more than a few cases. These figures, which give a fairly favorable picture of guidance for college women, should however be discounted for several reasons. They are based on returns from only a little more than half of the institutions interrogated and this half is probably the one in which most of the guidance is found. Furthermore the mere statement that an institution carries on some guidance

activity does not mean that it does so efficiently. The institutions that give guidance of a really serious nature are comparatively few. Some of them are described in detail.

An admirable feature of the book is a table showing the occupational distribution of the graduates of Mt. Holyoke College, which points to the moral that before one can give enlightened vocational guidance to college women he must know what are the occupations which these women enter.

One chapter deals with orientation classes for freshmen, another with extra-curricular activities. A chapter entitled Agencies in the Personnel Department treats psychological tests, rating scales, placement and personnel research.

With respect to the administration of guidance in a college, the author recommends the centralization of personnel activities in one department headed by a person who holds all responsibility.

> ". . . some one who can study the qualifications needed to profit by college training, one who can help to adjust the student in college, interpreting her to the college, and the college to her; some one who, with all this information at hand, can help to place the student after graduation, adjusting her as quickly as possible to occupational conditions. No one person can master all the knowledge which a personnel officer should possess. She should have associated with her specialists representing different fields of service, such as an appointment secretary, a vocational counselor, a psychiatrist, a psychologist, all cooperating and conferring frequently concerning the work of the department. The personnel worker should know the work of each sufficiently well to make the right contacts when the need arises. The Personnel Director should be a member of the faculty, with academic rank, an administrative officer with full authority in the supervision of the personnel office, but with close cooperation with other administrative officers, as well as the academic and student departments."

The author appends a complete set of forms on which to keep the records which should be maintained in a personnel office. Emphasis is laid on the need for using these records as a basis for research. A number of problems are cited which should be thus investigated.

H. D. Kitson.

Teachers College.

PSYCHOLOGICAL ABSTRACTS

The American Psychological Association began the publication of its new journal, *Psychological Abstracts*, with the January, 1927 number. It is edited by Professor Walter S. Hunter of Clark University,

assisted by the following distinguished cooperating editors from foreign countries: F. C. Bartlett, Cambridge University, V. M. Bekhterev, Leningrad, Ed. Claparede, University of Geneva, G. C. Ferrari, University of Bologna, A. Michotte, University of Louvain, H. Pieron, Sorbonne, M. L. Reymert, (formerly) University of Oslo, and W. Wirth, University of Leipzig.

The journal is international in character. It consists of non-critical abstracts of articles and books on psychological and cognate subjects, which appear as soon as possible after the original publication. These abstracts are prepared by competent psychologists in America and Europe. They are published in English, and average 150 words each.

The scope of *Psychological Abstracts* is indicated by the accompanying classification of its material: General Topics in Psychology, Sensation and Perception, Feeling and Emotion, Attention, Memory and Thought, Nervous System, Motor Phenomena and Action, Plant and Animal Behavior, Evolution and Heredity, Special Mental Conditions, Nervous and Mental Disorders, Social Functions of the Individual, Industrial and Personnel Problems, Childhood and Adolescence, Educational Psychology, Biometry and Statistics, and Mental Tests.

It is the intention of the Association to make *Psychological Abstracts* invaluable, not only to psychologists throughout the world, but also to all persons who are working in allied fields, such as physiology, psychiatry, education and biology.

Sample copies of the January and February issues will be sent for examination on request. The annual subscription rate is $6.00 (overseas $6.25), checks to be made payable to *Psychological Abstracts.* Subscriptions should be sent to H. S. Langfeld, Business Editor, Princeton University, Princeton, New Jersey.